Temas de matemáticas

Álgebra superior I

Segunda edición

Antonio Lascuraín Orive

las prensas de ciencias

Antonio Lascurain Orive

Álgebra superior I

Segunda edición

Facultad de Ciencias, UNAM
2014

Lascurain Orive, Antonio
Álgebra superior, 1 / Antonio Lascurain Orive - 2a. edición. - México : UNAM, Facultad de Ciencias, 2014.
x, 168 páginas : ilustraciones ; 22 cm. - (Temas de matemáticas) (Las prensas de ciencias)
Incluye índice
Bibliografía: página 163
ISBN 978-607-02-5458-1

1. Álgebra Estudio y enseñanza (Superior). 2. Álgebra. Problemas, ejercicios, etc. I. Universidad Nacional Autónoma de México. Facultad de Ciencias. II. Título. III. Serie. IV. Serie: Las prensas de ciencias.

512-scdd21 Biblioteca Nacional de México

Álgebra superior I.
1ª edición, 2 de agosto de 2012
2ª edición, 20 de mayo de 2014
1ª reimpresión, 2017

© D.R. 2014. Universidad Nacional Autónoma de México.
Facultad de Ciencias.
Ciudad Universitaria. Delegación Coyoacán,
C. P. 04510, México, Distrito Federal.
editoriales@ciencias.unam.mx

ISBN: 978-607-02-5458-1

Diseño de portada: Laura Uribe Hernández

Prohibida la reproducción parcial o total de la obra por cualquier medio,
sin la autorización por escrito del titular de los derechos patrimoniales.

Impreso y hecho en México.

A Adda Stella

A Guadalupe y Antonio

Prólogo

Este libro trata de algunos temas introductorios de Álgebra que se han enseñado en la Facultad de Ciencias de la UNAM en las últimas décadas, en el primer semestre de las carreras de Matemáticas, de Actuaría y de Física. Estos temas son fundamentales para todas las ramas de las matemáticas.

El objetivo es que los alumnos de los primeros semestres de las carreras de Matemáticas, Actuaría, Ciencias de la Computación y Física cuenten con un texto simple y breve donde puedan entender sin mucha dificultad los temas que se cubren en el curso Álgebra Superior I. Esta materia es un fundamento esencial en la formación de los estudiantes de estas carreras, y no solamente de aquéllos que se van a especializar en Álgebra. Los temas que se discuten son los del programa vigente, es decir, conjuntos, funciones, relaciones de equivalencia, inducción, cálculo combinatorio, el espacio vectorial $\mathbb{R}^n$, matrices, determinantes y sistemas de ecuaciones lineales. Además se incluye el tema del anillo de los enteros y los anillos $\mathbb{Z}_m$.

Existen dos buenos libros sobre el tema: el primero, *Álgebra superior*, de los autores H. Cárdenas, E. Lluis, F. Raggi y F. Tomás [2], y el segundo, editado en años recientes con el mismo título, de los autores A. Bravo, C. Rincón y H. Rincón [1]. Estos dos libros han constituido la herramienta de los estudiantes en los últimos años para esta importante materia del inicio de la carrera. Sin embargo, el primero [2] contiene un número excesivo de erratas, no ha sido actualizado y en algunos temas es impreciso y poco formal; algunos resultados no los prueba, como el hecho de que los anillos $\mathbb{Z}_m$, en efecto lo son. El segundo [1] es, en cierto sentido, más adecuado que el primero, al ser más reciente y al profundizar más en los temas; el problema radica en que es demasiado extenso.

En particular, el libro *Álgebra superior*, de Cárdenas *et al.* [2] es, de alguna manera, bastante bueno. Esto se debe, en parte, a que contiene ejemplos muy didácticos, como son los de la baraja inglesa. Otra cualidad que presenta, es

la de estar ordenado conforme al temario vigente. Sin embargo, desde que se editó, hace ya más de cuarenta años, sólo se han hecho reimpresiones (sin revisión), y no parece que se vaya a actualizar. Este hecho, junto con la demanda de mis estudiantes por mis notas manuscritas, motivó la elaboración del presente libro, el cual retoma varios de los ejemplos de [2], pero enmarca la teoría en un discurso matemático más actual. Cabe señalar que esta materia tiene un alto índice de reprobación, ya que un sector amplio de los estudiantes viene de la preparatoria con una formación deficiente, por lo que un texto de apoyo como el presente puede coadyuvar a mejorar el aprovechamiento de los alumnos.

El objetivo de este libro, basado principalmente en el de Cárdenas *et al* [2], es presentar los temas de manera simple y rigurosa. En la parte inicial se hace énfasis en algunos aspectos de la lógica formal, con el objeto de describir la simbología y facilitar el manejo de las pruebas formales en matemáticas. En la parte de combinatoria se hace claramente la diferencia de la parte formal y de la intuitiva. En general, se subraya la relación con otras ramas como el cálculo y la geometría analítica, por ejemplo, en el capítulo del espacio vectorial $\mathbb{R}^n$, se hace una breve descripción de las ecuaciones de los planos; en el capítulo de determinantes, se proporcionan también las ideas geométricas de este tema. Además, en el capítulo de ecuaciones, probando algunos teoremas simples de álgebra lineal, se demuestran de manera rigurosa los resultados necesarios para resolver cualquier sistema de m ecuaciones con n incógnitas; la inclusión de estos resultados, fáciles de probar, ciertamente aclaran al estudiante el tema, el cual se entiende mejor a la luz de la teoría que mediante una aplicación mecánica de algoritmos. Finalmente, en el capítulo de los enteros se prueba formalmente que los conjuntos $\mathbb{Z}_m$ son en efecto anillos.

En este libro aparecen también algunos temas más avanzados que no forman parte del material básico que se pretende cubrir en un curso diseñado para el primer semestre de la carrera, la razón de incluirlos es motivar a los estudiantes. Por ejemplo, la discusión del significado geométrico del determinante ciertamente es un tema más avanzado, por lo que se sugiere verlo al final del curso y no incluirlo como tema por examinar.

En resumen, el espíritu del libro es proporcionar a los alumnos un texto breve, simple, formal, que pone énfasis en que las matemáticas no son ramas aisladas, sino que interactúan unas con otras.

Los temas de este libro se pueden cubrir en un semestre, una posible distribución de ellos es la siguiente: cinco semanas para cubrir los primeros dos capítulos (conjuntos, funciones, inducción, relaciones de equivalencia y combinatoria); cinco semanas para estudiar el espacio vectorial $\mathbb{R}^n$, las matrices y las permutaciones; cinco semanas para los temas de los determinantes (aspectos algebraicos), las ecuaciones y los enteros, y la última semana, para la interpretación geométrica del determinante.

Agradezco a Cristina Angélica Núñez Rodríguez por la captura en Latex de mis notas para el curso Álgebra Superior I —el cual he impartido en múltiples ocasiones— así como por la elaboración de las figuras. Quiero agradecer también a Manuel Flores Galicia y a uno de los árbitros, quienes leyeron el texto de manera minuciosa.

Mi gratitud asimismo a los colegas que me han enriquecido con sus comentarios sobre la enseñanza de esta asignatura, así como a varios de mis alumnos por sus pertinentes intervenciones; a las autoridades de la Facultad de Ciencias y a la Dirección General de Asuntos del Personal Académico (DGAPA), que me apoyaron en la publicación de este libro con el proyecto PAPIME PE-103811. Finalmente, a la Coordinación de Servicios Editoriales, por su cuidadosa revisión, y a Héctor Cejudo Camacho por la corrección de algunas figuras del texto.

En esta segunda edición se hicieron diversas correcciones y se incluyeron nuevos ejercicios.

Contenido

1. Fundamentos **1**
1.1. Conjuntos 1
1.2. Subconjuntos 2
1.3. Operaciones con conjuntos 3
1.4. Producto cartesiano 8
1.5. Relaciones 9
1.6. Funciones 10
1.7. Composición de funciones 12
1.8. Funciones inyectivas, suprayectivas y biyectivas 15
1.9. Cardinalidad y conjuntos finitos 17
1.10. Inducción matemática 19
1.11. El teorema del binomio 21
1.12. Relaciones de equivalencia y particiones 23
1.13. Estructuras numéricas y algebraicas 28

2. Cálculo combinatorio **31**
2.1. Ordenaciones con repetición (versión intuitiva) 31
2.2. Ordenaciones (versión intuitiva) 32
2.3. Permutaciones (versión intuitiva) 33
2.4. Combinaciones 34
2.5. Funciones (2a visita) 36
2.6. Funciones inyectivas, suprayectivas y biyectivas (2a visita) 39
2.7. Ordenaciones con repetición (versión formal) 41
2.8. Ordenaciones (versión formal) 43
2.9. Permutaciones (versión formal) 45
2.10. Combinaciones y coeficientes binomiales 46

3. El espacio vectorial $\mathbb{R}^n$ 53
3.1. Vectores y sus operaciones 53
3.2. $\mathbb{R}^n$ 60
3.3. Subespacios vectoriales 63
3.4. Combinaciones lineales, dependencia e independencia lineal 69
3.5. Bases 74

4. Matrices y determinantes 81
4.1. Definiciones 81
4.2. El rango de una matriz 83
4.3. Permutaciones 88
4.4. Determinantes 93
4.5. Propiedades de los determinantes 95
4.6. Desarrollo por menores 102
4.7. Cálculo de determinantes 107
4.8. Caracterización del rango de una matriz usando determinantes 109
4.9. El determinante como área o volumen 113

5. Sistemas de ecuaciones lineales 121
5.1. Preliminares 121
5.2. Existencia de soluciones 123
5.3. Sistemas de n ecuaciones y n incógnitas 126
5.4. Sistemas homogéneos, funciones lineales 132
5.5. Sistema homogéneo asociado 136
5.6. Resolución de sistemas 138

6. Los anillos $\mathbb{Z}$ y $\mathbb{Z}_m$ 147
6.1. Anillos 147
6.2. Anillos $\mathbb{Z}_m$ 148
6.3. Propiedades de los enteros 152
6.4. Orden y unidades en $\mathbb{Z}$ 155
6.5. Principio de inducción 158

Glosario de símbolos 161

Bibliografía 163

Índice analítico 165

Capítulo 1

Fundamentos

1.1. Conjuntos

No se profundizará en la definición axiomática de conjunto, simplemente se tratará de manera intuitiva como una *colección* de *elementos*, por ejemplo, una colección de libros, de peces o de números. *Se dirá que 2 conjuntos son iguales si tienen los mismos elementos.*

Se usarán las letras mayúsculas $A, B, C...$ para los conjuntos y las minúsculas $a, b, c, \ldots, n, m, \ldots$ para los elementos. Para especificar los elementos de un conjunto se usarán llaves, por ejemplo

$$A = \{a, b, c\}.$$

Un conjunto importante son los enteros positivos llamados números naturales, denotado por $\mathbb{N} = \{1, 2, 3, ...\}$, y por supuesto los enteros

$$\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}.$$

Notación: $x \in A$ significa que el elemento x pertenece al conjunto A, y $x \notin A$ querrá decir que x no pertenece al conjunto A.

Ejemplos

1) Sea el conjunto $A = \{1, 3, 5, 7\}$. Se tiene que $5 \in A$ y $6 \notin A$.

2) Sea $A = \{1, 4, 9, 16, ..., n^2, ...\}$. En este caso $169 \in A$, pero $50 \notin A$.

3) El conjunto de las letras de la palabra Uaxactún es $\{a, c, n, t, u, x\}$.

Otros ejemplos importantes son los números reales, que son los puntos de la recta, este conjunto se denota por $\mathbb{R}$; por otro lado el plano cartesiano

$$\mathbb{R}^2 = \{(x, y) \,|\, x, y \in \mathbb{R}\},$$

así como sus subconjuntos, por ejemplo, la recta $y = 3x + 3$.

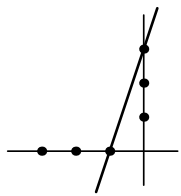


Figura 1.1: La recta $y = 3x + 3$ es un subconjunto del plano

El símbolo $\emptyset$ se usará para describir el conjunto que no tiene elementos; a este conjunto se le llama conjunto vacío.

Es conveniente usar condiciones para describir conjuntos:

$$\begin{aligned} \{2, 4, 6, 8, ...\} &= \{n \in \mathbb{N} \mid n \text{ es par}\} \\ &= \{n \in \mathbb{N} \mid n = 2m, m \in \mathbb{N}\}, \end{aligned}$$

o

$$\{1, 3, 5, 7, 9\} = \{n \in \mathbb{N} \mid 1 \leq n \leq 9 \quad \text{y} \quad n \text{ es impar}\}.$$

Otro ejemplo sería

$$\{1, 4, 9, 25, 36, ..., m^2, ...\} = \{n \in \mathbb{N} \mid n = m^2,\ m \in \mathbb{N}\}.$$

1.2. Subconjuntos

Definición 1 *Sean A y B conjuntos, se dice que B es un subconjunto de A, si cada elemento de B lo es también de A, se denota $B \subset A$, en caso contrario se escribirá $B \not\subset A$.*

Obsérvese que si $B \subset A$, se tiene $x \in B \Rightarrow x \in A$ y, viceversa, si para todo $x \in B$ se tiene $x \in A$, entonces $B \subset A$. En general, cuando la proposición P se cumple si y sólo si se cumple la proposición Q, escribiremos $P \iff Q$.

Bajo esta notación, la observación anterior se puede reescribir así

$$B \subset A \iff (x \in B \Rightarrow x \in A)\,.$$

Si $B \subset A$, se puede escribir también $A \supset B$, y se dirá que B está contenido en A, o que A contiene a B.

Ejemplos

1. Si $A = \{golondrinas\}$, $B = \{aves\}$ y $C = \{reptiles\}$, entonces $A \subset B$, pero $B \not\subset C$.
2. $A = \{n \in \mathbb{N} \mid 1 \leq n \leq 7\}$, $B = \{2, 3, 7\}$ y $C = \{2, 3, 8\}$, entonces $B \subset A$, pero $B \not\subset C$.
3. Si $z, w \in L$, donde L es una recta en $\mathbb{R}^2$, y si $A = \{t \in L \,|\, t \in \overline{zw}\}$, donde $\overline{zw}$ es el segmento en L que une a z con w, entonces $z \in L$ y $A \subset L$.

1.3. Operaciones con conjuntos

Al comparar dos conjuntos es conveniente pensar que ambos son subconjuntos de un mismo conjunto fijo, llamado universal.

Definición 2 *La unión de dos conjuntos A y B se define como el conjunto:*

$$A \cup B = \{x \mid x \in A \quad o \quad x \in B\}.$$

Las propiedades siguientes son consecuencia inmediata de la definición.

i) $A \subset A \cup B$, $B \subset A \cup B$,

ii) $A \cup B = B \cup A$ (conmutatividad),

iii) $(A \cup B) \cup C = A \cup (B \cup C)$ (asociatividad).

En virtud de la última observación, iii), al denotar la unión de más de dos conjuntos, no es necesario escribir los paréntesis.

Definición 3 *La intersección de dos conjuntos A y B se define como*

$$A \cap B = \{x \mid x \in A \ \ y \ \ x \in B\}.$$

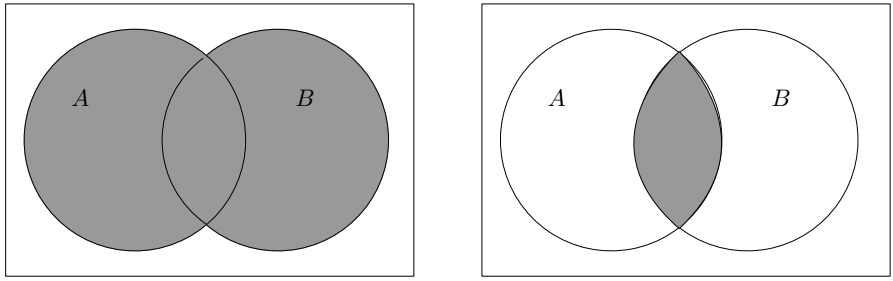
De nuevo se sigue de manera inmediata que

iv) $A \cap B \subset A, \ A \cap B \subset B,$

v) $A \cap B = B \cap A$ (conmutatividad),

vi) $(A \cap B) \cap C = A \cap (B \cap C)$ (asociatividad).

Como antes se pueden quitar los paréntesis.



$A \cup B$ $\qquad$ $A \cap B$

Figura 1.2: Unión e intersección de conjuntos

Proposición 1.3.1 *(Ley distributiva) Sean A, B y C conjuntos, entonces se tiene*

vii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

viii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$

Demostración. Probamos *vii)* y dejamos *viii)* como ejercicio.
Léase primero $\Rightarrow$ y luego $\Leftarrow$

$$\begin{aligned} x \in A \cap (B \cup C) &\Leftrightarrow x \in A \ \ y \ \ x \in (B \cup C) \\ &\Leftrightarrow x \in A \ \ y \ \ (x \in B \ \ o \ \ x \in C) \\ &\Leftrightarrow (x \in A \ \ y \ \ x \in B) \ \ o \ \ (x \in A \ \ y \ \ x \in C) \\ &\Leftrightarrow x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

□ [1]

[1] □ significa "fin de la prueba".

Otras técnicas de prueba muestran que

$$(A \cap B) \cup (A \cap C) \quad \subset \quad A \cap (B \cup C) :$$

Se tiene $A \cap B \subset A$ y $A \cap C \subset A$, además $A \cap B \subset B \subset B \cup C$ y $A \cap C \subset C \subset B \cup C$, por lo cual

$$(A \cap B) \subset A \cap (B \cup C) \quad y \quad (A \cap C) \subset A \cap (B \cup C).$$

Recordamos que un conjunto universal es un conjunto fijo que contiene a todos los conjuntos en discusión. Por ejemplo, en la geometría analítica el plano $\mathbb{R}^2$ es el conjunto universal y los subconjuntos son las rectas, las parábolas, los círculos, etcétera. En otras geometrías, los subconjuntos del plano pueden ser objetos simples o muy complejos como los *fractales*.

Definición 4 *Sea $\mathbb{X}$ un conjunto universal y A un subconjunto de $\mathbb{X}$. Se define el complemento del conjunto A, denotado por A^c, como los elementos de $\mathbb{X}$ que no pertenecen a A, específicamente*

$$A^c = \{x \,|\, x \in \mathbb{X}, x \notin A\}.$$

Evidentemente el complemento de un conjunto varía si el universal donde vive cambia, por ejemplo, el complemento de $A = \{1, 2\}$ en $\mathbb{X} = \{1, 2, 3\}$ es $\{3\}$ pero en $\mathbb{X} = \{1, 2, 3, 4\}$ es $\{3, 4\}$.

Las propiedades básicas del complemento son:

$ix)$ $(A^c)^c = A$,

$x)$ $A \cup A^c = \mathbb{X}$,

$xi)$ $A \cap A^c = \emptyset$.

Las propiedades $x)$ y $xi)$ se siguen directamente de la definición. Para probar $ix)$ tómese $x \in (A^c)^c$, entonces como $x \in \mathbb{X}$ y $x \notin A^c$, se sigue de $x)$ que $x \in A$. Viceversa, si $x \in A$, entonces $x \in \mathbb{X}$ pero $x \notin A^c$ $\therefore x \in (A^c)^c$. [2]

[2] $\therefore$ significa "de donde".

Otra propiedad importante es

xii) $A \subset B \Leftrightarrow B^c \subset A^c$.

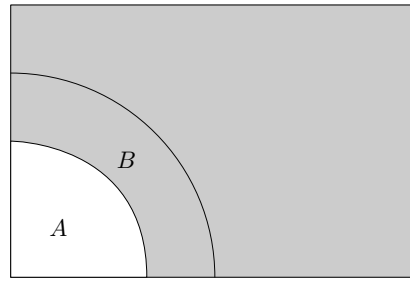


Figura 1.3: La contención de conjuntos se invierte al tomar complementos

Demostración.
$\Rightarrow$) Se debe probar que si $x \in B^c$, entonces $x \in A^c$. Suponiendo que no es cierto lo afirmado se tendría que existe una $x \in B^c$ que no pertenece a A^c, *i.e.*, [3] $x \notin A^c$, por lo que $x \in A$, pero entonces se sigue por hipótesis que $x \in B$, lo cual contradice $x \in B^c$. Por lo tanto se concluye que suponer lo contrario de lo afirmado llevó a una contradicción, por lo que lo afirmado debe ser cierto.[4]

$\Leftarrow$) Esta propiedad es dual de la anterior, sólo tendríamos que escribir $E = B^c$ y $F = A^c$. Por la primera parte, como $E \subset F$, se sigue $F^c \subset E^c$, *i.e.*, $A \subset B$. □

Proposición 1.3.2 *(Leyes de De Morgan) Sean A y B conjuntos, entonces*

xiii) $(A \cup B)^c = A^c \cap B^c$,

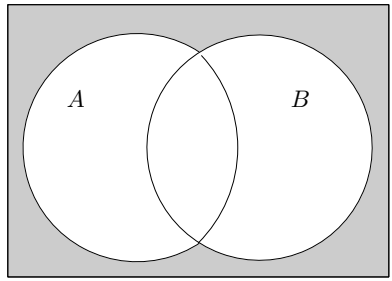
xiv) $(A \cap B)^c = A^c \cup B^c$.

[3] *i.e.*, del latín *id est*, "es decir".

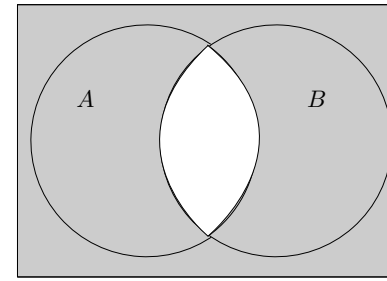
[4] Este tipo de demostración, usada por Aristóteles, se le llama por reducción al absurdo, y conlleva el razonamiento elemental de la lógica formal que dice:

(La proposición $M \Rightarrow$ La proposición N) $\Leftrightarrow$ $(\sim N \Rightarrow \sim M)$,

donde el símbolo $\sim$ significa: no.



$(A \cup B)^c = A^c \cap B^c$



$(A \cap B)^c = A^c \cup B^c$

Figura 1.4: Complementos de unión e intersección

DEMOSTRACIÓN. Probamos *xiv)* y dejamos *xiii)* como ejercicio.
Léase primero $\Rightarrow$ y luego $\Leftarrow$

$$\begin{aligned} x \in (A \cap B)^c &\Leftrightarrow x \notin (A \cap B) \\ &\Leftrightarrow x \notin A \quad o \quad x \notin B \\ &\Leftrightarrow x \in A^c \quad o \quad x \in B^c \\ &\Leftrightarrow x \in A^c \cup B^c. \end{aligned}$$

□

Otra demostración de

$$(A \cap B)^c \supset A^c \cup B^c :$$

Se tiene $A \cap B \subset A$ y $A \cap B \subset B$, por lo que $(A \cap B)^c \supset A^c$ y $(A \cap B)^c \supset B^c$, entonces

$$(A \cap B)^c \supset A^c \cup B^c.$$

Definición 5 *La diferencia entre dos conjuntos A y B es el conjunto:*

$$A - B = \{x \mid x \in A \quad y \quad x \notin B\}.$$

Nótese que $A - B = A \cap B^c$.

Proposición 1.3.3 $A - (B \cap C) = (A - B) \cup (A - C)$.

Demostración.

$$\begin{aligned} A-(B\cap C) &= A\cap (B\cap C)^c \\ &= A\cap (B^c\cup C^c) \\ &= (A\cap B^c)\cup (A\cap C^c) \\ &= (A-B)\cup (A-C). \end{aligned}$$

□

EJERCICIOS 1.3
1. Demuestre la Proposición 1.3.1 *viii*).
2. Demuestre la Proposición 1.3.2 *xiii*).

1.4. Producto cartesiano

Definición 6 *Sea A un conjunto, las parejas ordenadas de A son aquéllas de la forma (a,b), tal que $a\in A$ y $b\in A$.*

Nótese, que dadas dos parejas ordenadas (a,b), y (c,d), éstas son iguales si y sólo si $a=c$ y $b=d$. Por ejemplo, si $A=\{1,2\}$ las parejas ordenadas de A son $(1,1),(1,2),(2,1),(2,2)$. Otro ejemplo es el plano cartesiano que consiste de las parejas ordenadas de reales:

$$\mathbb{R}\times\mathbb{R}=\{(x,y)\,|\,x,y\in\mathbb{R}\}.$$

También se pueden considerar parejas ordenadas de un par de conjuntos A y B, de tal manera que el primer elemento esté en A y el segundo en B.

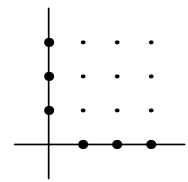


Figura 1.5: Retícula $\mathbb{N}\times\mathbb{N}$

Definición 7 *Sean A y B conjuntos, el producto cartesiano de A y B, denotado por $A\times B$, es el conjunto de parejas ordenadas*

$$A\times B=\{(a,b)\mid a\in A \quad y \quad b\in B\}.$$

Por ejemplo, si $A = \{1, 2\}$ y $B = \{a, b\}$, entonces

$$A \times B = \{(1, a), (1, b), (2, a), (2, b)\}.$$

Otros ejemplo es la retícula descrita en la Figura 1.5

$$\mathbb{N} \times \mathbb{N} = \{(n, m) \mid n \in \mathbb{N} \quad \text{y} \quad m \in \mathbb{N}\}.$$

También son ejemplos de productos cartesianos $\mathbb{Z} \times \mathbb{Z}$ y $\mathbb{R} \times \mathbb{R}$.

Obsérvese que $(a, b) = (c, d)$ si y sólo si $a = c$ y $b = d$. En particular, $(1, 2) \neq (2, 1)$ ($\neq$ significa es distinto). Se escribirá A^2 por $A \times A$.

Dado un conjunto A se define otro llamado *conjunto potencia* cuyos elementos son los subconjuntos de A. Por ejemplo, si $A = \{1, 2\}$, su conjunto potencia es $\{\{1\}, \{2\}, \{1, 2\}, \emptyset\}$. Por convención el vacío es subconjunto de todo conjunto A, ya que $\forall\, x \in \emptyset$, se tiene $x \in A$. [5] Nótese que el conjunto potencia también se puede pensar como una *familia* de conjuntos.

1.5. Relaciones

Definición 8 *Sean A y B conjuntos, una relación entre A y B es un subconjunto del producto cartesiano.*

Por ejemplo, si $A = \{a\}$ y $B = \{1, 2\}$, $A \times B$ tiene 2 relaciones de un elemento $R_1 = \{(a, 1)\}$ y $R_2 = \{(a, 2)\}$, también tiene una sola relación de 2 elementos (la total) $R_3 = \{(a, 1), (a, 2)\}$, finalmente contiene la relación vacía, *i.e.*, $\emptyset$, por lo que $A \times B$ tiene exactamente 4 relaciones.

Encontramos ahora las relaciones entre A y B, donde $A = \{a, b\}$ y $B = \{1, 2\}$. Como $A \times B = \{(a, 1), (a, 2), (b, 1), (b, 2)\}$, éstas son:

$$\begin{array}{ll}
R_1 = \{(a, 1)\}, & R_9 = \{(a, 2), (b, 2)\}, \\
R_2 = \{(a, 2)\}, & R_{10} = \{(b, 1), (b, 2)\}, \\
R_3 = \{(b, 1)\}, & R_{11} = \{(a, 1), (a, 2), (b, 1)\}, \\
R_4 = \{(b, 2)\}, & R_{12} = \{(a, 1), (a, 2), (b, 2)\}, \\
R_5 = \{(a, 1), (a, 2)\}, & R_{13} = \{(a, 1), (b, 1), (b, 2)\}, \\
R_6 = \{(a, 1), (b, 1)\}, & R_{14} = \{(a, 2), (b, 1), (b, 2)\}, \\
R_7 = \{(a, 1), (b, 2)\}, & R_{15} = \{(a, 1), (a, 2), (b, 1), (b, 2)\}, \\
R_8 = \{(a, 2), (b, 1)\}, & R_{16} = \emptyset.
\end{array}$$

[5] $\forall$ significa "para todo".

Definición 9 *El dominio de una relación* $R \subset A \times B$ *se define como:*

$$D_R = \{a \in A \mid \exists\, b \in B,\ (a,b) \in R\}.^{6}$$

En los ejemplos anteriores se tiene

$$D_{R_5} = \{a\}, \quad D_{R_7} = A, \quad D_{R_3} = \{b\}.$$

Definición 10 *La imagen de una relación* $R \subset A \times B$ *se define como:*

$$I_R = \{b \in B \mid \exists\, a \in A, \quad \textit{tal que} \quad (a,b) \in R\}.$$

Usando los ejemplos anteriores se tendría

$$I_{R_3} = \{1\}, \quad I_{R_5} = B, \quad I_{R_2} = \{2\}.$$

El codominio de una relación $A \times B$ es el segundo factor B, obsérvese que la imagen de una relación siempre es un subconjunto del codominio.

1.6. Funciones

Una subclase de las relaciones son particularmente importantes.

Definición 11 *Sea* R *una relación entre* A *y* B. *Se dice que* R *es una función si* R *cumple las siguientes 2 condiciones:*

1) $D_R = A$, *i.e.,* $\forall\, x \in A \quad \exists\, b \in B \ \ni\ (a,b) \in R.$ [7]

2) Cada elemento en A *tiene asociado un único elemento en* B, *i.e., si* $(x,y_1) \in R$ *y* $(x,y_2) \in R$, *entonces necesariamente* $y_1 = y_2$.

Si $R \subset A \times B$ es una función, la pareja $(x,y) \in R$ se escribirá como $(x, f(x))$ y la función se denotará por

$$f : A \longrightarrow B \quad \text{o} \quad A \xrightarrow{f} B$$

y se dirá que $f(x)$ es la imagen de x bajo f. Al conjunto A se le llama dominio de la función y a B el codominio.

[6] $\exists$ significa "existe".
[7] $\ni$ significa "tal que".

Ejemplos

1. Retomando los ejemplos de la sección anterior $A = \{a, b\}$ y $B = \{1, 2\}$, se tiene que $R_8 = \{(a, 2), (b, 1)\}$ es una función, sin embargo $R_2 = \{(a, 2)\}$ y $R_{10} = \{(b, 1), (b, 2)\}$ no lo son, ¿por qué?

2. $f : \mathbb{N} \longrightarrow \mathbb{N}$ dada por la regla de correspondencia $f(n) = n^2 + 1$ es una función.

3. Sean $A = \{seres\, humanos\}$, $B = \{países\}$, $f : A \longrightarrow B$ la relación que a cada persona le asocia su país de nacimiento, y $g : A \longrightarrow B$ la relación que a cada persona le asocia su nacionalidad. Entonces, f es función pero g no lo es, ¿por qué?

4. Sea A cualquier conjunto, la relación $I_A : A \longrightarrow A$ que tiene la regla de correspondencia $I_A(x) = x \quad \forall\, x \in A$, llamada función idéntica, es una función.

Obsérvese que 2 funciones

$$f : A \longrightarrow B \quad \text{y} \quad g : C \longrightarrow D$$

son iguales si y sólo si

a) $A = C$,

b) $B = D$,

c) $f(x) = g(x) \quad \forall\, x \in A$.

Esto nos permite reformular nuestra definición: *Una función* $f : A \longrightarrow B$ *es una regla de correspondencia que a cada elemento del conjunto* A *le asocia uno y sólo un elemento de* B.

En resumen, una función consiste de 3 cosas: 2 conjuntos (dominio y codominio) y una regla de correspondencia.

Definición 12 *La imagen de una función* $f : A \longrightarrow B$ *es*

$$Im\, f = \{b \in B \mid \exists\, a \in A,\ f(a) = b\}.$$

Ejemplos

1. Sean $A = \{1,2\}$, $B = \{a,b\}$, $f : A \to B$ dada por $f(1) = a$, y $f(2) = a$, entonces $Im\ f = \{a\}$.

2. Sea $f : \mathbb{N} \cup \{0\} \longrightarrow \mathbb{N}$ dada por $f(n) = n+1$, entonces $Im\ f = \mathbb{N}$.

Obsérvese que $Im\ f$ siempre es un subconjunto del codominio. Trabajando con funciones cuyo codominio es un conjunto finito, y donde no es necesario especificar el codominio, algunas veces conviene denotarlas de la siguiente manera

$$f = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix},$$

donde $A = \{a_1, a_2, ..., a_n\}$, $b_i \in B \quad \forall i$, $f : A \longrightarrow B$ y $f(a_i) = b_i \quad \forall i$.

Por ejemplo, si $A = \{1,2,3,4,5\}$, entonces $f(x) = x^2$ se puede denotar

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{pmatrix},$$

el codominio puede ser $\mathbb{N}$ o $\mathbb{Z}$ o cualquier conjunto finito que contenga a $\{1,4,9,16,25\}$.

EJERCICIO 1.6
1. Sean $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ y $g : \mathbb{N} \longrightarrow \mathbb{Z}$ dadas por $f(n) = n^2 + 1 = g(n)$, ¿son iguales?

1.7. Composición de funciones

Si el codominio de una función coincide con el dominio de otra se puede construir una nueva función.

Definición 13 *Sean $f : A \longrightarrow B$ y $g : B \longrightarrow C$ funciones, se puede construir entonces una nueva función $g \circ f : A \longrightarrow C$ con dominio A, codominio C y con regla de correspondencia*

$$x \longrightarrow g(f(x)),$$

A esta función se le llama f compuesta con (o seguida de) g, y se denota por $g \circ f$. También se escribe $g \circ f(x) = g(f(x))$.

Ejemplos

1. Sean $f : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $f(x) = x^2 + 1$ y $g : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $g(x) = 4x - 1$, entonces $g \circ f : \mathbb{R} \longrightarrow \mathbb{R}$ y $f \circ g : \mathbb{R} \longrightarrow \mathbb{R}$ están dadas por

$$g \circ f(x) = g(x^2 + 1) = 4(x^2 + 1) - 1 = 4x^2 + 3,$$

$$\text{y} \quad f \circ g(x) = f(4x - 1) = (4x - 1)^2 + 1 = 16x^2 - 8x + 2.$$

2. Sean $A = \{a_1, a_2, a_3\}$, $B = \{b_1, b_2\}$, $C = \{c_1, c_2, c_3\}$ y

$$f = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_1 & b_2 \end{pmatrix} \quad y \quad g = \begin{pmatrix} b_1 & b_2 \\ c_1 & c_2 \end{pmatrix},$$

entonces

$$g \circ f = \begin{pmatrix} a_1 & a_2 & a_3 \\ c_1 & c_1 & c_2 \end{pmatrix}.$$

3. Sean $A = \{1, 2\}$ y $f : A \longrightarrow A$ dada por

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

entonces $f^2 = f \circ f = I_A$.

4. Sean $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ dadas por $f(x) = x^3$ y $g(x) = -2$, entonces

$$f \circ g(x) = f(-2) = (-2)^3 = -8 \quad \text{y} \quad g \circ f(x) = g(x^3) = -2.$$

Obsérvese que si $f : A \longrightarrow B$ es una función, entonces

$$I_B \circ f = f \quad \text{y} \quad f \circ I_A = f.$$

Proposición 1.7.1 *Sean $f : A \longrightarrow B$, $g : B \longrightarrow C$ y $h : C \longrightarrow D$ funciones, entonces*

$$h \circ (g \circ f) = (h \circ g) \circ f,$$

es decir, la composición de funciones es asociativa y por lo tanto se pueden omitir los paréntesis.

Demostración. Como ambas composiciones tienen dominio A y codominio D, basta verificar que tienen la misma regla de correspondencia.

Si $x \in A$,

$$h \circ (g \circ f)(x) = h[g \circ f(x)] = h[g(f(x))],$$

también

$$(h \circ g) \circ f(x) = h \circ g(f(x)) = h[g(f(x))].$$

□

Definición 14 *Sea $f : A \longrightarrow B$ una función, si existe otra función de la forma $g : B \longrightarrow A \ni g \circ f = I_A$, a g se le llama inversa izquierda de f, (a f inversa derecha de g).*

Por ejemplo, sea $g : \mathbb{Z} \longrightarrow \mathbb{Z}$ definida por $g(n) = [n/2]$, [8] y sea $f : \mathbb{Z} \longrightarrow \mathbb{Z}$, dada por $f(n) = 2n$, entonces

$$g \circ f(n) = g(2n) = [2n/2] = [n] = n \quad \text{y} \quad g \circ f = I_{\mathbb{Z}},$$

i.e., g es inversa izquierda de f.

De manera análoga se define la inversa derecha de $f : A \longrightarrow B$, como otra función $h : B \longrightarrow A \ni f \circ h = I_B$. Obsérvese que en el ejemplo anterior la función g no tiene a f como inversa izquierda, ya que, por ejemplo:

$$f \circ g(3) = f([3/2]) = f(1) = 2 \quad i.e., \quad f \circ g \neq I_{\mathbb{Z}}.$$

Sin embargo, como se probó, f es inversa derecha de g (ya que g es inversa izquierda de f).

Definición 15 *Sean $f : A \longrightarrow B$ y $g : B \longrightarrow A \ni g \circ f = I_A$, y $f \circ g = I_B$, entonces se dice que f es invertible y a g se le llama simplemente inversa.*

Por ejemplo, si $f : \mathbb{R} \longrightarrow \mathbb{R}$, está dada por $f(x) = 2x$, y $g : \mathbb{R} \longrightarrow \mathbb{R}$, es tal que $g(x) = x/2$, entonces g es inversa de f.

Teorema 1.7.2 *Si f tiene inversa izquierda g_1 e inversa derecha g_2, entonces $g_1 = g_2$, i.e., f es invertible.*

[8] La función $[x]$ es el máximo entero menor o igual a x, por ejemplo $[2.1] = 2$, $[0.3] = 0$, $[-1.2] = -2$.

DEMOSTRACIÓN. Por hipótesis si $f : A \longrightarrow B$ se tiene $g_1 \circ f = I_A$ y $f \circ g_2 = I_B$, por lo que

$$g_1 = g_1 \circ I_B = g_1 \circ f \circ g_2 = I_A \circ g_2 = g_2.$$

□

Corolario 1.7.3 *Sea $f : A \longrightarrow B$ invertible, entonces su inverso $g : B \longrightarrow A$ es único.*

EJERCICIO 1.7
1. Sean $f : \mathbb{R} \longrightarrow \mathbb{R}$, $g : \mathbb{R} \longrightarrow \mathbb{R}$ dadas por $f(x) = 4x^2 + x + 1$ y $g(x) = 2x - 3$, calcule $g \circ f$ y $f \circ g$.

1.8. Funciones inyectivas, suprayectivas y biyectivas

Definición 16 *Una función $f : A \longrightarrow B$ se llama inyectiva si $\forall x_1, x_2 \in A$, tales que $x_1 \neq x_2$, se tiene $f(x_1) \neq f(x_2)$.*

Ejemplos

1. $f(x) = x + 1, \quad f : \mathbb{R} \longrightarrow \mathbb{R}.$

2. $f(n) = 2n, \qquad f : \mathbb{Z} \longrightarrow \mathbb{Z}.$

Es claro que estas funciones son inyectivas ya que mandan puntos distintos en puntos distintos, sin embargo la función $f : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $f(x) = x^2$ no es inyectiva, ya que $f(1) = f(-1) = 1$.

En general, si la afirmación M dice $x_1, x_2 \in A, x_1 \neq x_2$ y la afirmación N dice $f(x_1) \neq f(x_2)$, probar que f es inyectiva es mostrar que $M \Rightarrow N$, para esto (como ya se mencionó) basta demostrar que $\sim N \Rightarrow \sim M$, esto es, f es inyectiva, si dados $a_1, a_2 \in A$, tales que $f(a_1) = f(a_2)$ se debe tener $a_1 = a_2$.

Por ejemplo, $f(x) = 2x + 3$ es inyectiva, ya que si $2x + 3 = 2y + 3$, entonces, $2x = 2y$ y $x = y$ (ésta es la forma más común de probar que una función es inyectiva).

Definición 17 *Sea* $f : A \longrightarrow B$, *se dice que* f *es suprayectiva (o sobre) si* $Imf = B$, *i.e., si* $\forall\, y \in B\ \exists\, x \in A\ \ni\ f(x) = y$.

Retomando los ejemplos anteriores, $f : \mathbb{R} \longrightarrow \mathbb{R}$, dada por $f(x) = x+1$ es suprayectiva, ya que si $y \in \mathbb{R}$, escribiendo $y = x + 1$, se puede tomar $x = y - 1$ y $f(y - 1) = y$.

Sin embargo, $f : \mathbb{Z} \longrightarrow \mathbb{Z}$, dada por $f(n) = 2n$ no es suprayectiva, ya que la imagen de esta función consiste sólo de los números pares.

Definición 18 *A una función que es inyectiva y suprayectiva se le llama biyectiva.*

La función $f : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $f(x) = 3x + 1$ es biyectiva. Primero, ésta es inyectiva ya que si $f(x) = f(y)$, entonces $3x + 1 = 3y + 1$ $\therefore\ 3x = 3y$ y $x = y$. También es suprayectiva, ya que dada $y \in \mathbb{R}$, escribiendo $y = 3x + 1$ y despejando se tiene $x = \frac{y-1}{3}$. Por lo cual, $f(\frac{y-1}{3}) = 3\,(\frac{y-1}{3}) + 1 = y$.

Como último ejemplo se toma la función $g : \mathbb{Z} \longrightarrow \mathbb{Z}$ dada por

$$g(n) = \left[\frac{n}{2}\right],$$

g es suprayectiva ya que dada $m \in \mathbb{Z}$, tomando $x = 2m$,

$$g(2m) = \left[\frac{2m}{2}\right] = m,$$

sin embargo, g no es inyectiva, ya que

$$g(2n+1) = \left[\frac{2n+1}{2}\right] = n = \left[\frac{2n}{2}\right] = g(2n).$$

Proposición 1.8.1 *Una función* $f : A \longrightarrow B$ *es biyectiva si y sólo si es invertible.*

Demostración. Si f es biyectiva se define $g : B \longrightarrow A$ de la manera obvia: a $y \in B$ se le asocia la única $x \in A$ que cumple $f(x) = y$ (la existencia de dicha x se sigue de la inyectividad y la suprayectividad). Escribiendo $g(y) = x$ se tiene que si $y \in B$ $f \circ g(y) = f(x) = y$ y $f \circ g = I_B$, y si $x \in A$ $g \circ f(x) = g(y) = x$ por lo que $g \circ f = I_A$.

Viceversa, si f es invertible sea $g : B \longrightarrow A$ su inversa. En este caso, si $f(a_1) = f(a_2)$, entonces $g \circ f(a_1) = g \circ f(a_2)$ y $a_1 = a_2$, por lo cual f es inyectiva. Asimismo, si $y \in B$, tomando $x \in A$ de tal manera que $g(y) = x$, se tiene que $y = f \circ g(y) = f(x)$ y f es suprayectiva.

□

Obsérvese que la composición de funciones inyectivas es inyectiva: dadas $g : A \longrightarrow B$ y $f : B \longrightarrow C$ funciones inyectivas, $a_1, a_2 \in A$ tales que $f \circ g(a_1) = f \circ g(a_2)$ se sigue que $g(a_1) = g(a_2)$, ya que f es inyectiva, más aún $a_1 = a_2$ por la inyectividad de g.

También si $f : A \longrightarrow B$ y $g : B \longrightarrow C$ son suprayectivas se tiene que $g \circ f$ lo es, esto se sigue ya que si $c \in C$ existe $b \in B \ni g(b) = c$ y también $a \in A \ni f(a) = b$ (por la suprayectividad de f y g) $\therefore g \circ f(a) = c$ y por consiguiente $g \circ f$ es suprayectiva.

EJERCICIOS 1.8
1. Sea $f(x) = x^2 + 1$ $f : \mathbb{R} \longrightarrow \mathbb{R}$ pruebe que f no es inyectiva ni tampoco suprayectiva.
2. Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $f(x) = 4x + 7$, pruebe que f es biyectiva; encuentre la inversa.
3. Exhiba $f : \mathbb{N} \longrightarrow \mathbb{N}$ y $g : \mathbb{N} \longrightarrow \mathbb{N}$ tales que f es inyectiva, g suprayectiva, pero $g \circ f$ no sea inyectiva ni suprayectiva.

1.9. Cardinalidad y conjuntos finitos

Definición 19 *Dados dos conjuntos A y B se dice que tienen la misma cardinalidad si existe $f : A \longrightarrow B$ biyectiva.*

Se usará el símbolo I_n para denotar el conjunto de los primeros n naturales,

$$I_n = \{1, 2, 3, ..., n\} = \{a \in \mathbb{N} \mid 1 \leq a \leq n\}.$$

Definición 20 *Un conjunto A es finito si existe una biyección (función biyectiva)*

$$f : A \longrightarrow I_n,$$

para alguna $n \in \mathbb{N}$. En este caso se dice que el cardinal de A, denotado por $\#A$, es n.

Definición 21 *Sea A un conjunto que no es finito, entonces se dice que A es infinito.*

Con lo anterior se tienen las siguientes observaciones que relacionan las funciones y la cardinalidad de conjuntos finitos.

1. Sea $f : A \longrightarrow B$ una función inyectiva entre dos conjuntos finitos, entonces
$$\#A \leq \#B.$$
Para probar esto, nótese que si $A = \{a_1, a_2, ..., a_n\}$, entonces el conjunto $\{f(a_1), f(a_2), ..., f(a_n)\}$ consiste de elementos distintos (ya que f es inyectiva).

2. Sea $f : A \longrightarrow B$ una función suprayectiva entre dos conjuntos finitos, entonces
$$\#A \geq \#B.$$
Esto se cumple, ya que si $B = \{b_1, b_2, ..., b_m\}$, como f es suprayectiva existen $a_1, a_2, ..., a_n \in A$ tales que $f(a_i) = b_i$, $i = 1, 2, ..., m$, y además las a_i son todas distintas, ya que f es función.

3. Sea $f : A \longrightarrow B$ una función biyectiva entre dos conjuntos finitos, entonces
$$\#A = \#B.$$
Por las dos observaciones anteriores se da la igualdad.

Proposición 1.9.1 *Sea $f : A \longrightarrow B$ una función, donde A y B son finitos con la misma cardinalidad, entonces,*

$$f \text{ es inyectiva} \Leftrightarrow f \text{ es suprayectiva.}$$

Demostración. Se tiene como hipótesis que: $\#A = n = \#B$.
$\Rightarrow$) Dado que f es inyectiva, se tienen n imágenes distintas, *i.e.*, todos los elementos de B están en la imagen.
$\Leftarrow$) Si f no es inyectiva, entonces $\#Imf < n$, pero eso contradice la hipótesis.

□

Esta propiedad no es válida para conjuntos infinitos, por ejemplo, la función $f : \mathbb{N} \longrightarrow \mathbb{N}$, $f(n) = 2n$ es inyectiva pero no suprayectiva. También, como ya se probó, la función $g : \mathbb{Z} \longrightarrow \mathbb{Z}$, dada por $g(n) = [n/2]$, es suprayectiva pero no inyectiva.

1.10. Inducción matemática

La inducción matemática consiste en una manera de probar múltiples afirmaciones o propiedades (tantas como los naturales), y que dependen de los mismos números naturales. Por ejemplo si se quiere probar que

$$1+2+\cdots+n \;=\; \frac{n(n+1)}{2}, \tag{1.1}$$

obsérvese que cada $n \in \mathbb{N}$ define una igualdad, el *axioma o principio de inducción* establece que si se cumplen dos condiciones, la propiedad es válida para todos los naturales, estas condiciones son:

A) 1 cumple la propiedad,

B) si la propiedad es válida para n, entonces también lo es para $n+1$.

En nuestro ejemplo particular probaremos que se cumplen las dos condiciones y concluiremos que (1.1) se cumple $\forall\, n \in \mathbb{N}$:

A) $1 = \frac{1(2)}{2}$ por lo que (1.1) se cumple si $n = 1$,

B) si (1.1) es válida para n, entonces

$$1+2+\cdots+n \;=\; \frac{n(n+1)}{2},$$

sumando $n+1$ se tiene

$$\begin{aligned} 1+2+\cdots+n+n+1 \;&=\; \frac{n(n+1)}{2} + \; n+1 \\ = \; (n+1)\left(\frac{n}{2}+1\right) \;&=\; \frac{(n+1)(n+2)}{2}, \end{aligned}$$

es decir, (1.1) se cumple para $n+1$ y se sigue del principio de inducción que (1.1) se cumple $\forall\, n \in \mathbb{N}$.

Intuitivamente es clara la validez del principio: si vale para 1, vale para 2, si también es válido para 2, entonces también para 3, etcétera.

Ejemplos

1) Se prueba que

$$n^3 - n \quad \text{es un múltiplo de } 6 \quad \forall\, n \in \mathbb{N}. \tag{1.2}$$

Para probar (1.2) se aplica inducción probando A y B :

A) si $n = 1 \Rightarrow 1^3 - 1 = 6 \cdot 0$,

B) si se cumple (1.2) para m, se tiene $m^3 - m = 6 \cdot k, \quad k \in \mathbb{Z}$. Ahora

$$\begin{aligned}(m+1)^3 - (m+1) &= m^3 + 3m^2 + 3m + 1 - m - 1 \\ &= m^3 - m + 3(m^2 + m) \\ &= 6k + 6t, \quad t \in \mathbb{Z},\end{aligned}$$

i.e., (1.2) se cumple para $m+1$ y por lo tanto $\forall\, m \in \mathbb{Z}$ (usamos el hecho de que m^2+m siempre es par: $(2n+1)^2 = 4n^2+4n+1$).

2) Se afirma que

$$2^n < n! \quad \text{si} \quad n \geq 4, \ ^9 \tag{1.3}$$

Nótese que basta probar, usando inducción, que

$$2^{n+3} < (n+3)! \quad \forall\, n \in \mathbb{N}.$$

A) $2^{1+3} = 16 < (1+3)! = 24$,

B) si $2^{n+3} < (n+3)!$, como $2 < n+4$ se tiene
$2 \cdot 2^{n+3} < (n+3)!(n+4)$, esto es $2^{n+4} < (n+4)!$
$\therefore$ (1.3) se cumple $\forall\, n \geq 4$.

EJERCICIOS 1.10

1. Demuestre por inducción que si $r \in \mathbb{R}$, $r \neq 1$

$$1 + r + r^2 + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r}.$$

2. Usando inducción pruebe que

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

3. Demuestre que $1 + 3 + 5 + 7 + \cdots + 2n - 1 = n^2$.

4. Usando inducción muestre que

$$2 + 5 + 8 + \cdots + (3n-1) = \frac{n\,(3n+1)}{2}.$$

[9] El símbolo $n!$ se le llama n factorial, se define como $1 \cdot 2 \cdot 3 \cdot \cdots (n-1)(n)$, por ejemplo, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$.

1.11. El teorema del binomio

Recordamos de la sección anterior que dado un número natural n, n-factorial, denotado por $n!$, está dado por $n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$.

Por convención $0! = 1$. Obsérvese que se puede definir $n!$ inductivamente, esto es, $n! = n\,(n-1)!$. Por ejemplo, $4! = 4\,(3)! = 24$.

Teorema 1.11.1 *(Pascal) Sean n,m enteros no negativos y*

$$C_n^m = \frac{n!}{m!(n-m)!}, \quad n \geq m,$$

entonces se tiene

$$C_{n-1}^m + C_{n-1}^{m-1} = C_n^m.$$

DEMOSTRACIÓN. Sumamos

$$\frac{(n-1)!}{m!(n-1-m)!} + \frac{(n-1)!}{(m-1)!(n-1-(m-1))!}$$

$$= \frac{(n-1)!}{m!(n-m-1)!} + \frac{(n-1)!}{(m-1)!(n-m)!} = \frac{(n-1)!(n-m)}{m!(n-m)!} + \frac{(n-1)!m}{m!(n-m)!}$$

$$= \frac{(n-1)![(n-m)+m]}{m!(n-m)!} = \frac{n!}{m!(n-m)!} = C_n^m.$$

□

Corolario 1.11.2 $C_n^m \in \mathbb{N}$

DEMOSTRACIÓN. Lo probamos por inducción sobre n.

A) Si $n = 1$ $\quad C_1^1 = \frac{1!}{1!(1-1)!} = 1 \quad$ y $\quad C_1^0 = \frac{1!}{0!(1-0)!} = 1,$

B) Si suponemos cierto el teorema para $n-1$ se tiene que tanto C_{n-1}^m como C_{n-1}^{m-1} son números naturales, y por lo tanto en virtud del teorema de Pascal, se sigue que C_n^m también es un número natural □

Teorema 1.11.3 *(Del binomio de Newton) Si $a, b \in \mathbb{R}$ y n es un entero no negativo, entonces*

$$(a+b)^n = C_n^0 a^n + C_n^1 a^{n-1}b + \cdots + C_n^n b^n.$$

Por ejemplo: $(a+b)^2 = a^2+2ab+b^2$ y $(a+b)^3 = a^3+3a^2b+3ab^2+b^3$, ya que $C_2^1 = \frac{2!}{1!(2-1)!}$, $C_3^2 = \frac{3!}{2!(3-2)!}$, etcétera. Obsérvese que los superíndices y los subíndices que definen el teorema crecen y decrecen conforme a la siguiente fórmula [10]

$$\sum_{j=0}^{n} C_n^j \, a^{n-j} \, b^j.$$

Demostración. Probamos el teorema usando inducción sobre n.

A) Si $n = 1$ $\quad (a+b)^1 = C_1^0 \, a^1 + C_1^1 \, a^0 \, b^1 = 1 \cdot a + 1 \cdot b = a + b.$

B) Si el teorema es cierto para $n-1$, se tiene

$$(a+b)^{n-1} = \sum_{j=0}^{n-1} C_{n-1}^j \, a^{n-1-j} \, b^j,$$

y entonces

$$\begin{aligned}(a+b)^n = (a+b)(a+b)^{n-1} &= a\sum_{j=0}^{n-1} C_{n-1}^j a^{n-1-j} b^j + b\sum_{j=0}^{n-1} C_{n-1}^j a^{n-1-j} b^j \\ &= \sum_{j=0}^{n-1} C_{n-1}^j a^{n-j} b^j + \sum_{j=0}^{n-1} C_{n-1}^j a^{n-1-j} b^{j+1},\end{aligned}$$

expresando la segunda sumatoria en términos de $k = j+1$, en lugar de j, *i.e.*, corriendo los índices, se tiene

$$(a+b)^n = \sum_{j=0}^{n-1} C_{n-1}^j \, a^{n-j} \, b^j + \sum_{k=1}^{n} C_{n-1}^{k-1} \, a^{n-k} \, b^k.$$

[10] $\sum$ significa "sumatoria", y cada índice desde $j = 0$ hasta n representa un sumando, *i.e.*, el teorema del binomio dice:

$$(a+b)^n = \sum_{j=0}^{n} C_n^j \, a^{n-j} \, b^j.$$

Finalmente, homologando índices y factorizando

$$\begin{aligned}(a+b)^n &= \sum_{i=1}^{n-1} C_{n-1}^{i} a^{n-i}b^i + \sum_{i=1}^{n-1} C_{n-1}^{i-1} a^{n-i}b^i + C_{n-1}^{0}a^n + C_{n-1}^{n-1}b^n \\ &= \sum_{i=1}^{n-1} (C_{n-1}^{i} + C_{n-1}^{i-1}) a^{n-i}b^i + C_n^0 a^n + C_n^n b^n \\ &= \sum_{i=0}^{n} C_n^i a^{n-i}b^i,\end{aligned}$$

en virtud del teorema de Pascal. □

1.12. Relaciones de equivalencia y particiones

Definición 22 *Una relación $R \subset A \times A$ se llama de equivalencia si satisface:*

1) $(a,a) \in R \qquad \forall\, a \in A \qquad$ *(reflexividad),*

2) si $(a,b) \in R$, *entonces* $(b,a) \in R \qquad$ *(simetría),*

3) si $(a,b) \in R$ *y* $(b,c) \in R$, *entonces* $(a,c) \in R \quad$ *(transitividad).*

Obsérvese que la simetría se puede expresar como:

$$(a,b) \in R \Leftrightarrow (b,a) \in R.$$

Ejemplos

1. Sea A cualquier conjunto, entonces la mínima relación de equivalencia es la diagonal
$$R = \{(a,a) \,|\, a \in A\},$$
i.e., toda relación de equivalencia incluye a la diagonal, ¿por qué?

2. Sea A la familia de los triángulos del plano y
$$R = \{(a,b) \in A \times A \,|\, a,\ b \text{ son semejantes}\},$$
evidentemente R es de equivalencia.

3. Sea $A = \{1, 2, 3\}$ y

$$R = \{(1, 1), (1, 2), (2, 1)(2, 2)\},$$

entonces R no es de equivalencia. Aunque es simétrica y transitiva, no es reflexiva pues $(3, 3) \notin R$.

4. Sea $A = \{1, 2, 3\}$ y

$$R = \{(1, 1), (2, 2), (3, 3)(1, 2), (2, 1), (2, 3), (3, 2)\},$$

R tampoco es de equivalencia, ya que no es transitiva pues $(1, 3) \notin R$.

Dada cualquier relación de equivalencia R, escribiremos

$$a \sim b \quad si \quad (a, b) \in R.$$

Definición 23 *Sea A un conjunto, una partición P de A consiste de una familia de subconjuntos de A no vacíos $\{A_i\}_{i \in I}$, tales que*

1) si $A_i \neq A_j$, entonces $A_i \cap A_j = \emptyset$,

2) $\bigcup_{i \in I} A_i = A$.

Ejemplos

1. Si $A = \{1, 2, 3, 4, 5\}$, una partición de A es

$$A_1 = \{1, 2\}, \quad A_2 = \{3\} \quad \text{y} \quad A_3 = \{4, 5\}.$$

2. Si $A = \mathbb{N}$, una partición de A es

$$A_1 = \{\text{pares}\} \quad \text{y} \quad A_2 = \{\text{nones}\}.$$

3. Sea $A = \{\text{instrumentos de una orquesta sinfónica}\}$ y

$$A_1 = \{\text{cuerdas}\}, \quad A_2 = \{\text{alientos}\}, \quad A_3 = \{\text{metales}\},$$

$$A_4 = \{\text{percusiones}\}.$$

Entonces estos subconjuntos no forman una partición de A, ya que el clarinete es un aliento y también es un metal.

4. Si $A = \{1, 2, 3, .., 10\}$ y

$$A_1 = \{1, 3, 5\}, \quad A_2 = \{1, 4, 6\} \quad \text{y} \quad A_3 = \{2, 7, 8, 9, 10\},$$

entonces A_1, A_2, A_3 no forman una partición de A pues $1 \in A_1 \cap A_2$.

Teorema 1.12.1 *Sea A cualquier conjunto y R una relación de equivalencia en $A \times A$, entonces R induce una partición en A.*

Demostración. Sea $x \in A$, se define

$$A_x = \{y \in A \,|\, y \sim x\},$$

donde $y \sim x$ significa $(y, x) \in R$. Obsérvese que $A_x \neq \emptyset$, ya que $x \in A_x$. Se afirma que $\{A_x\}_{x \in A}$ es una partición P de A.

La propiedad 2) es inmediata, ya que como $A_x \subset A$, $\forall\, x \in A$, ciertamente $\bigcup_{x \in A} A_x \subset A$. Además, como dada $x \in A$, se tiene que $x \in A_x$ se sigue que $A \subset \bigcup_{x \in A} A_x$.

Para probar la propiedad 1) supongamos que $z \in A_x \cap A_y$, donde $x, y, z \in A$, se debe probar que $A_x = A_y$.

Como $z \in A_x \cap A_y$, se sigue que $z \sim x$ y $z \sim y$, por lo que $x \sim y$ (simetría y transitividad).

Finalmente, probamos $A_x \subset A_y$ (y viceversa). Sea $w \in A_x$, entonces $w \sim x$ y por transitividad $w \sim y$, *i.e.*, $w \in A_y \therefore A_x \subset A_y$. La otra contención se prueba de la misma manera.

□

El recíproco también es cierto.

Teorema 1.12.2 *Sea A un conjunto arbitrario y P una partición en A, entonces P induce una relación de equivalencia en A.*

Demostración. Se define una relación R en $A \times A$ de la siguiente manera $(x, y) \in R$ si existe A_i en $P \ni x, y \in A_i$. Probamos que esta relación es de equivalencia.

Evidentemente $x \sim x \quad \forall\, x$. También es obvio que si $x \sim y$ entonces $y \sim x$. Finalmente, si $x \sim y$ y $y \sim z$, entonces existe A_i y A_j en la partición tales que $x, y \in A_i$ y $y, z \in A_j$, como A_i y A_j se intersectan deben ser iguales *i.e.*, $x, z \in A_i = A_j$, y $x \sim z$ (por lo que la relación es transitiva).

□

Definición 24 *Dada una relación de equivalencia R en $A \times A$, a los subconjuntos de la partición inducida por R se les llama clases de equivalencia.*

Ejemplos

1. Sean A un conjunto arbitrario y R la relación diagonal, *i.e.*,

$$R = \{(x, x) \,|\, x \in A\},$$

entonces la partición inducida por R es

$$A = \bigcup_{x \in A} \{x\}.$$

2. Sean $A = \{1, 2, 3\}$ y $R = \{(1,1), (1,2), (2,1), (2,2), (3,3)\}$, entonces la partición inducida por R es

$$A = \{1, 2\} \cup \{3\}.$$

3. Sea $m \in \mathbb{N}$ fijo. Se define una relación $R \subset \mathbb{Z} \times \mathbb{Z}$ de la siguiente manera: $(x, y) \in R$ si $x - y = mt, \quad t \in \mathbb{Z}$. Esta relación es de equivalencia (ejercicio), y la partición inducida es

$$\mathbb{Z} = \bigcup_{j=0}^{m-1} [j],$$

donde

$$[j] = \{s \in \mathbb{Z} \,|\, s - j = mt, \quad t \in \mathbb{Z}\}, \quad j \in \{0, 1, 2, ..., m-1\}.$$

Esto se sigue del algoritmo de la división y del hecho de que dados $h \neq k, \ h, k \in \mathbb{N} \cup \{0\} \ni h, k < m$, entonces m no es un factor de $h - k$. El algoritmo de la división establece que dado $t \in \mathbb{Z}$ y $m \in \mathbb{N}$ existen s, r únicos tales que

$$t = sm + r, \qquad 0 \leq r < m.$$

En el caso en que $t < 0$, tomando $-t = sm + r$

$$t = -sm - r - m + m = -(s+1)m + (m - r),$$

$(m - r)$ es el nuevo residuo, cf. [2] o [5].

Por ejemplo, si $m = 3$ la partición es

$$\mathbb{Z} = A_0 \cup A_1 \cup A_2 \qquad \text{donde}$$
$$A_j = \{3t + j \,|\, t \in \mathbb{Z}\} \qquad j = 0, 1, 2.$$

Sea R_A el conjunto de las relaciones de equivalencia en un conjunto A y P_A las particiones de A. Los teoremas anteriores muestran que existen funciones

$$\varphi : R_A \longrightarrow P_A \qquad y \qquad \psi : P_A \longrightarrow R_A,$$

definidas por la asociación natural.

Proposición 1.12.3 *Sean φ, ψ, R_A, P_A y A, como se indicó arriba, entonces se tiene*

a) $\psi \circ \varphi = I_{R_A}$,

b) $\varphi \circ \psi = I_{P_A}$.

DEMOSTRACIÓN.

a) Dada una relación de equivalencia R en A, $\varphi(R)$ es la partición definida por las clases de equivalencia de A determinadas por R y ciertamente $\psi\varphi(R) = R$, ya que la definición de ψ establece que dos elementos de A están relacionados si pertenecen al mismo subconjunto (estos subconjuntos son las clases de equivalencia de R).

b) Dada P una partición de A, $\psi(P)$ es una relación de equivalencia cuyas clases de equivalencia son precisamente los subconjuntos de P y se concluye que $\varphi\psi(P)$ es de nuevo P. Esto se sigue, ya que por definición φ le asocia a $\psi(P)$ la partición definida por sus clases de equivalencia.

□

EJERCICIOS 1.12

1. Demuestre que la siguiente relación es de equivalencia. Sea $m \in \mathbb{N}$ fijo, se define una relación $R \subset \mathbb{Z} \times \mathbb{Z}$ de la siguiente manera: $(x, y) \in R$ si $x - y = mt$, $t \in \mathbb{Z}$.

2. Describa la partición en $\mathbb{Z}$ inducida por la relación de equivalencia descrita en el Ejercicio 1. Pruebe sus afirmaciones.

3. Defina en $\mathbb{R}$, $x \sim y$, si $x - y \in \mathbb{Z}$, pruebe que esta relación es de equivalencia y demuestre que cada clase de equivalencia tiene un único representante en el intervalo $[0, 1)$.

1.13. Estructuras numéricas y algebraicas

Las estructuras algebraicas más elementales son las numéricas, es decir, conjuntos de números que cumplen ciertos axiomas. Algunas de éstas son:

Los naturales

$$\mathbb{N} = \{1, 2, 3, ...\};$$

el anillo de los números enteros

$$\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\};$$

el campo de los números racionales, denotados por $\mathbb{Q}$,

$$\mathbb{Q} = \left\{ \frac{a}{b} \,|\, a, b, \in \mathbb{Z}, \quad b \neq 0 \right\};$$

el campo de los números reales

$$\mathbb{R} = \{A.b_1 b_2 b_3 ...\},$$

donde $A \in \mathbb{Z}$ y $b_j \in \{0, 1, , 2, ..., 9\}$, *i.e*, los números enteros con expansiones decimales infinitas, y sin colas infinitas de nueves; estos puntos se pueden pensar como los puntos de la recta;

el campo de los números complejos

$$\mathbb{C} = \{a + ib \,|\, a, b \in \mathbb{R}\},$$

estos puntos pueden pensarse como los puntos del plano, donde $i = (0, 1)$ satisface $i^2 = -1$,

Una operación binaria en un conjunto A es una función

$$\mu : A \times A \longrightarrow A.$$

Por ejemplo, en los enteros la suma y el producto son operaciones binarias que satisfacen ciertas propiedades.

SUMA

S1) $a + b = b + a \quad \forall\, a, b \in \mathbb{Z}$ (conmutatividad),

S2) $(a + b) + c = a + (b + c) \quad \forall\, a, b, c \in \mathbb{Z}$ (asociatividad),

S3) $\exists\, 0 \in \mathbb{Z} \ni a + 0 = a \quad \forall\, a \in \mathbb{Z}$ (existencia del neutro aditivo),

S4) $a + (-a) = 0 \quad \forall\, a \in \mathbb{Z}$ (existencia del inverso aditivo).

PRODUCTO

P1) $a\,b = b\,a \quad \forall\, a, b \in \mathbb{Z}$ (conmutatividad),

P2) $(a\,b)\,c = a\,(b\,c) \quad \forall\, a, b, c \in \mathbb{Z}$ (asociatividad),

P3) $\exists\, 1 \in \mathbb{Z} \ni a \cdot 1 = a \quad \forall\, a \in \mathbb{Z}$ (existencia del neutro multiplicativo),

DISTRIBUTIVIDAD

D) $a\,(b + c) = a\,b + a\,c \quad \forall\, a, b, c \in \mathbb{Z}.$

Estas propiedades implican otra.

LEY DE LA CANCELACIÓN DE LA SUMA

$$Si \quad a + c = b + c, \quad entonces \quad a = b.$$

Esta ley se cumple, ya que $(a + c) + (-c) = (b + c) + (-c) \Rightarrow a = b.$

Los números racionales, reales y complejos también cumplen estas propiedades y otra más:

P4) $a\,a^{-1} = 1 \quad \forall\, a \neq 0$ (existencia del inverso multiplicativo).

De manera análoga a la suma se tiene la siguiente propiedad.

LEY DE LA CANCELACIÓN DEL PRODUCTO

$$Si \quad a\,b = a\,c, \quad a \neq 0, \quad entonces \quad b = c.$$

Esto se sigue, al suponer que $a\,b = a\,c$, ya que entonces se tiene que $a^{-1}(a\,b) = a^{-1}(a\,c)$ y $b = c$.

En estas estructuras (salvo en $\mathbb{C}$) existe también una relación de orden, denotada por $<$, que satisface las siguientes cuatro propiedades:

a) si $a < b$ y $b < c$, entonces $a < c$ (transitividad),

b) dados a y b se cumple una y sólo una de las siguientes afirmaciones:

$i)$ $a = b$,
$ii)$ $a < b$,
$iii)$ $b < a$,

a esta propiedad se le llama ley de tricotomía,

c) si $a < b \quad \Rightarrow \quad a + c < b + c \qquad \forall\, a, b, c$
(compatibilidad del orden con la suma),

d) si $a < b \quad$ y $\quad 0 < c, \quad$ entonces $ac < bc \qquad \forall\, a, b, c$
(compatibilidad del orden con el producto).

Una estructura algebraica que satisface $S1), ..., S4), P1), P2), P3)$ *y* $D)$ se le llama *anillo conmutativo con unidad*, si además cumple $P4)$ se le llama *campo*.

Si un anillo cumple también las propiedades de orden se le llama *anillo ordenado*, por ejemplo $\mathbb{Z}$. Ejemplos de campos ordenados son $\mathbb{Q}$ y $\mathbb{R}$.

A la estructura algebraica más simple se le llama *grupo*; es un conjunto G con una operación binaria, que se le llama usualmente *producto*, de tal manera que cada elemento tiene un inverso, existe un neutro y el *producto* es asociativo. Por ejemplo, el conjunto de biyecciones (funciones biyectivas) de un conjunto cualquiera en sí mismo constituye un grupo, donde el *producto* es la composición, léase de derecha a izquierda.

EJERCICIOS 1.13

1. Demuestre la última afirmación de esta sección.

2. Exhiba una relación en un conjunto infinito que sea simétrica y transitiva pero que no sea reflexiva.

3. Muestre en un conjunto finito una relación que sea reflexiva y transitiva pero que no sea simétrica.

4. Exhiba una relación que sea reflexiva y simétrica pero que no sea transitiva.

5. Defina la unión e intersección de conjuntos para un número finito o infinito de éstos.

6. Generalice y pruebe las Proposiciones 1.3.1 y 1.3.2 para más de tres o dos conjuntos (un número finito de éstos).

Capítulo 2

Cálculo combinatorio

Describiremos primero los conceptos del cálculo combinatorio de manera intuitiva y posteriormente los estudiaremos rigurosamente.

2.1. Ordenaciones con repetición (versión intuitiva)

Consideremos el conjunto $A = \{a, b, c\}$. Al conjunto de 'palabras' de dos letras formadas con elementos de A se les llama ordenaciones con repetición de las letras a, b, c tomadas de 2 en 2 (decimos 'palabras' ya que solamente se trata de yuxtaponer letras). Se les llama con repetición ya que, por ejemplo, puede aparecer aa. En este caso hay nueve ordenaciones con repetición, y un método para obtener todas se ilustra en la Figura 2.1.

	a	b	c
a	aa	ab	ac
b	ba	bb	bc
c	ca	cb	cc

Figura 2.1: Tabla de ordenaciones con repetición de dos elementos

Si tomamos ahora como ejemplo las trasmisiones telegráficas de dos sonidos: uno corto denotado por $\cdot$ y uno largo denotado por $-$, a las señales de un sonido, se les puede considerar como las ordenaciones con repetición del conjunto $A = \{\cdot, -\}$ tomadas de uno en uno, éstas son:

$$\cdot \quad \text{y} \quad -$$

Las señales de dos sonidos serían las ordenaciones con repetición tomadas de 2 en 2, éstas las podemos derivar de las de un sonido, como se muestra en la siguiente tabla, el número de ellas es $2\times 2 = 4$.

	—	·
—	— —	· —
·	— ·	· ·

Figura 2.2: Tabla de ordenaciones con repetición de dos sonidos

Las de tres sonidos serían las ordenaciones con repetición tomadas de tres en tres, que a su vez se pueden derivar de la tabla anterior, el número de ellas es $2\times 4 = 8$. Son todas, ya que cualquier señal de tres sonidos consiste de una de un sonido seguida de una de dos sonidos.

	— —	— ·	· —	· ·
—	— — —	— — ·	— · —	— · ·
·	· — —	· — ·	· · —	· · ·

Figura 2.3: Tabla de ordenaciones con repetición de tres sonidos

Se denotará por OR_n^m al número de ordenaciones con repetición de un conjunto de n elementos tomados de m en m. Se deduce de los ejemplos anteriores que $OR_3^2 = 3^2$, $OR_2^1 = 2$, $OR_2^2 = 2^2$, $OR_2^3 = 2^3$. Probaremos posteriormente que $OR_n^m = n^m$.

2.2. Ordenaciones (versión intuitiva)

Supongamos ahora que se requiere formar palabras de dos letras tomadas del conjunto $\{a, b, c\}$, pero con la condición que éstas sean distintas, se tienen seis casos

$$ab,\ ac,\ ba,\ bc,\ ca,\ cb.$$

A estas palabras se les llamará ordenaciones del conjunto $\{a, b, c\}$, tomadas de 2 en 2.

Si consideramos ahora el conjunto $A = \{a, b, c, d\}$, las ordenaciones de A tomadas de 2 en 2 son 12:

$$ab,\ ac,\ ad,\ ba,\ bc,\ bd,\ ca,\ cb,\ cd,\ da,\ db,\ dc.$$

Obsérvese que el número de ordenaciones es el número de ordenaciones con repetición menos el número de ordenaciones con letras repetidas.

En nuestro ejemplo, el número de las ordenaciones es:

$$4^2 - 4 = 4(4-1) = 12.$$

Se denotará por O_n^m al número de ordenaciones en un conjunto de n elementos tomadas de m en m. Obsérvese que $m \leq n$, ¿por qué? Se sigue de nuestros ejemplos que

$$O_3^2 = 6, \quad O_4^2 = 12.$$

Se probará que $O_n^m = \frac{n!}{(n-m)!}$.

Consideremos ahora el siguiente problema, se tienen cuatro lienzos de colores: rojo, azul, verde y blanco. ¿Cuántas banderas tricolores se pueden formar? La pregunta es equivalente a encontrar las ordenaciones de 3 en 3 del conjunto $\{r, a, v, b\}$.

Un primer grupo de banderas sería,

$$rav,\ rab,\ rva,\ rvb,\ rba,\ rbv,$$

intercambiando el rojo con los otros tres colores se obtienen todas las posibilidades, *i.e.*, hay 24 banderas tricolores, lo cual coincide con nuestra fórmula

$$O_4^3 = \frac{4!}{(4-3)!} = 24.$$

2.3. Permutaciones (versión intuitiva)

A las ordenaciones de un conjunto de n elementos tomadas de n en n se les llama permutaciones de n elementos.

Por ejemplo: si $A = \{1, 2, 3\}$, las permutaciones de A son

$$123,\ 132,\ 213,\ 231,\ 312,\ 321.$$

Las banderas bicolores también se pueden pensar como permutaciones de dos elementos, cuando se trabaja con sólo dos colores.

La fórmula que mostramos para el número de ordenaciones exhibe como caso particular el número de permutaciones de un conjunto de n elementos. Este número, denotado por P_n, está dado por

$$P_n = O_n^n = \frac{n!}{(n-n)!} = n!.$$

Nótese que esta fórmula se aplica a nuestros ejemplos.

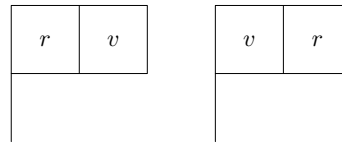


Figura 2.4: Banderas bicolores con dos lienzos verde y rojo

Como un último ejemplo calcularemos cuántos números de cuatro cifras distintas se pueden formar con los digitos 1, 3, 7 y 8.

Los que empiezan con 1 son:

$$1378,\ 1387,\ 1738,\ 1783,\ 1837 \text{ y } 1873,$$

intercambiando 1 con los otros tres dígitos se obtienen un total de 24, *i.e.*, 4!.

2.4. Combinaciones

Definición 25 *Sea A un conjunto de n elementos, a los subconjuntos de A que tienen m elementos se les llama combinaciones de los n elementos de A tomadas de m en m.*

Ejemplos

1. Sea $A = \{1, 2, 3\}$, las combinaciones de A tomadas de 2 en 2 son:

$$\{1, 2\},\ \{1, 3\} \text{ y } \{2, 3\},$$

tomadas de 1 en 1 son:

$$\{1\},\ \{2\} \text{ y } \{3\},$$

tomadas de 3 en 3 son:

$$\{1, 2, 3\}.$$

2. Si se tiene un grupo de cinco músicos y se quiere elegir un trío, el número posible de tríos es el número de combinaciones de un conjunto de cinco elementos tomados de 3 en 3; si los músicos son Andrea, Itzel, Laura, Pedro y Juan, los tríos posibles son:

$$\{A, I, L\},\ \{A, I, P\},\ \{A, I, J\},\ \{A, L, P\},\ \{A, L, J\},$$
$$\{A, P, J\},\ \{I, L, P\},\ \{I, L, J\},\ \{I, P, J\},\ \{L, P, J\},$$

por lo que se pueden formar 10 tríos.

3. Si ahora se quiere elegir un trío entre Andrea, Itzel, Laura y Pedro las posibilidades son :

$$\{A, I, L\},\ \{A, I, P\},\ \{A, L, P\},\ \text{y}\ \{I, L, P\},$$

éstas son cuatro, que es el mismo número de posibilidades para elegir una persona entre 4, *i.e.*, elegir 3 de 4 es eliminar 1 de 4.

Se probará que el número de combinaciones de un conjunto de n elementos tomadas de m en m está dado por los coeficientes binomiales (del teorema del binomio)

$$C_n^m = \frac{n!}{m!(n-m)!}. \tag{2.1}$$

Suponiendo cierto este hecho, usamos también el símbolo C_n^m para denotar el número de estas combinaciones.

Habiendo probado la fórmula de las ordenaciones se demostrará que

$$C_n^m P_m = O_n^m,$$

lo cual implica la fórmula (2.1), y justifica el uso del símbolo C_n^m como número de combinaciones, ya que

$$C_n^m m! = \frac{n!}{(n-m)!}.$$

La baraja inglesa provee de buenos ejemplos al análisis combinatorio, se tienen 52 cartas, cada una con un número y un palo, los números son $A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q$ y K, los palos son: {♦, ♥, ♣ y ♠}, es decir, diamantes, corazones, tréboles y espadas (o pics), respectivamente.

Dado que una mano de poker consta de cinco cartas, el número total de manos es

$$C_{52}^5 = \frac{52!}{5!47!} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2} = 52 \cdot 51 \cdot 10 \cdot 49 \cdot 2.$$

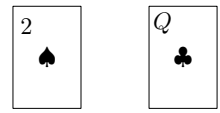


Figura 2.5: Ejemplos de cartas de baraja inglesa

2.5. Funciones (2a visita)

Las funciones surgen en una infinidad de situaciones, veamos unos ejemplos.

1. Sea A el conjunto de alumnos de la primaria "Benito Juárez" y B los salones de dicha primaria, la asociación de que a cada alumno le corresponde su salón es una función (en la Facultad de Ciencias esta asociación no sería una función).

2. Sea A el conjunto de casas de la calle Uxmal y $B = \mathbb{N}$, asignar a cada casa un número oficial es una función. Las casas que tienen dos números no serían ejemplo de función (por ejemplo, # 20 antes 5).

3. Las parejas ordenadas de números se pueden pensar también como funciones, *i.e.*, si $A = \{1\text{er lugar}, 2\text{o lugar}\}$, $B = \mathbb{Z}$, entonces $(2,5)$ o $(7,7)$ son funciones:

   $1\text{er lugar} \longrightarrow 2$,

   $2\text{o lugar} \longrightarrow 5$, etcétera.

4. Sea A un conjunto de 12 jugadores de beisbol (numerados del 1 al 12) y B el conjunto de posiciones en este deporte: lanzador, receptor, primera, segunda y tercera bases, jardineros izquierdo, derecho y central, esto es,

   $$B = \{L, R, P_b, S_b, T_b, J_I, J_D, J_C\}.$$

   Si el capitán asigna a cada posición un jugador, entonces está formando una función, digamos

   $$\begin{array}{lll} L \to 7 & S_b \to 2 & J_D \to 4 \\ R \to 11 & T_b \to 6 & J_C \to 10. \\ P_b \to 1 & J_I \to 5 & \end{array}$$

5. La función $f : \mathbb{Z} \longrightarrow \mathbb{Z}$, dada por $f(n) = n^3$.

En los ejemplos anteriores siempre se mencionó:

a) un conjunto A (el dominio) de: alumnos, casas, 1o y 2o lugares, posiciones de beisbol y los números enteros $\mathbb{Z}$;

b) otro conjunto B (el codominio) al cual se le asocian los elementos de A: salones, números naturales, números enteros, jugadores y de nuevo números enteros;

c) una forma de asociación en la que a cada elemento de A se le asigna un único elemento de B. A ésta se le llama *regla de correspondencia.*

Como se mencionó, estas funciones se designan por

$$f : A \longrightarrow B,$$

y $f(a) = b$, si $b \in B$ es el elemento de B asociado a a. Por ejemplo, $f(2) = 8$, en el Ejemplo 5.

Otros ejemplos

1. Sean $A = \{x, y, z\}$ y $B = \{1, -1\}$, la función $f : A \longrightarrow B$ dada por $f(x) = 1$, $f(y) = -1$, $f(z) = 1$, se puede expresar como

$$f = \begin{pmatrix} x & y & z \\ 1 & -1 & 1 \end{pmatrix}.$$

Una expresión de la forma $\begin{pmatrix} x & x & z \\ 1 & -1 & 1 \end{pmatrix}$ no sería función.

2. Un problema que puede ser relevante es saber cuántas funciones hay de un conjunto en otro. Un ejemplo trivial pero didáctico es el siguiente:

¿Cuántas banderas bicolores se pueden formar con tres colores: verde, blanco y rojo? Esta pregunta se puede expresar en términos de funciones como sigue. Sean $A = \{1, 2\}$ y $B = \{v, b, r\}$, donde 1 se refiere el lugar izquierdo de la bandera y 2 al derecho. De esta manera las banderas son funciones.

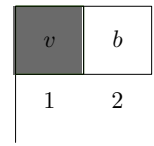


Figura 2.6: Bandera bicolor

Por ejemplo, la bandera en la Figura 2.6 está representada por la función

$$\begin{pmatrix} 1 & 2 \\ v & b \end{pmatrix},$$

las demás son

$$\begin{pmatrix} 1 & 2 \\ v & r \end{pmatrix} \begin{pmatrix} 1 & 2 \\ b & v \end{pmatrix} \begin{pmatrix} 1 & 2 \\ b & r \end{pmatrix} \begin{pmatrix} 1 & 2 \\ r & v \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ r & b \end{pmatrix} \begin{pmatrix} 1 & 2 \\ v & v \end{pmatrix} \begin{pmatrix} 1 & 2 \\ b & b \end{pmatrix} \begin{pmatrix} 1 & 2 \\ r & r \end{pmatrix},$$

sin embargo las últimas tres, aunque son funciones, no representan banderas bicolores, *i.e.*, hay seis banderas bicolores.

EJERCICIOS 2.5

1. Interprete como función el resultado de un examen en un grupo de 20 alumnos.
2. ¿Qué diagramas son funciones en la Figura 2.7?

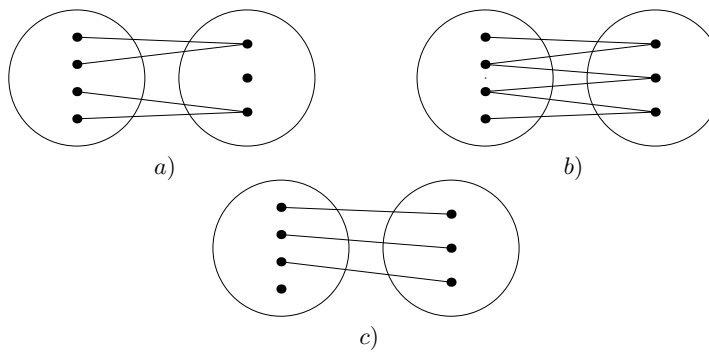


Figura 2.7: ¿Cuáles son funciones?

2.6. Funciones inyectivas, suprayectivas y biyectivas (2a visita)

Recordamos las definiciones. Sea $f : A \longrightarrow B$ una función.

a) Se dice que f es inyectiva si $\forall\ a_1, a_2 \in A\ \ni\ a_1 \neq a_2$, se tiene $f(a_1) \neq f(a_2)$.

b) Se dice que f es suprayectiva si $\forall\ y \in B$, existe $x \in A\ \ni\ f(x) = y$.

c) Se dice que f es biyectiva si es inyectiva y suprayectiva.

Ejemplos

1. El problema de buscar banderas bicolores con los colores $\{v, b, r\}$ se puede reformular y se afirma que hay seis al ser éste el número de funciones inyectivas que hay del conjunto $A = \{1, 2\}$ en $B = \{v, b, r\}$. Obsérvese que de las nueve funciones establecidas en dicho ejemplo ninguna es suprayectiva, ¿por qué?

2. Considérese los conjuntos $A = \{x, y, z\}$ y $B = \{7, 8\}$. Nótese que hay ocho funciones de A en B :

$$\begin{pmatrix} x & y & z \\ 7 & 7 & 7 \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ 7 & 7 & 8 \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ 7 & 8 & 7 \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ 7 & 8 & 8 \end{pmatrix}$$

$$\begin{pmatrix} x & y & z \\ 8 & 7 & 7 \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ 8 & 7 & 8 \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ 8 & 8 & 7 \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ 8 & 8 & 8 \end{pmatrix}.$$

Obsérvese que este procedimiento es el mismo que el de encontrar las ordenaciones con repetición de un conjunto de dos elementos tomados de 3 en 3, se tenía $OR_2^3 = 2^3$, ninguna de estas funciones es inyectiva, y todas, salvo la primera y la última, son suprayectivas.

En general, si $f : A \longrightarrow B$, A tiene n elementos y B m elementos, y $n > m$, entonces f no puede ser inyectiva, ¿por qué?

Obsérvese que una función es:

a) inyectiva, si a elementos distintos del dominio le corresponden elementos distintos del codominio,

b) suprayectiva, si su imagen es el codominio,

c) biyectiva, si a cada elemento del codominio le corresponde uno y sólo un elemento en el dominio.

Definición 26 *Sea* $f : A \longrightarrow B$ *una función, el subconjunto de* B

$$\{y \in B \,|\, \exists\, x \in A \ni f(x) = y\},$$

se llama la imagen de f *y se denota por* Imf.

Ejemplo: sea $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = 2x + 1$, f es inyectiva ya que si $f(x_1) = f(x_2)$, se tiene $2x_1 + 1 = 2x_2 + 1$ y $x_1 = x_2$. También es suprayectiva, ya que si $y \in \mathbb{R}$ $f(x) = 2x + 1 = y$ si y sólo si $x = \frac{y-1}{2}$, *i.e.*, $f(\frac{y-1}{2}) = 2(\frac{y-1}{2}) + 1 = y$.

Nótese que si se tienen dos funciones $g : A \longrightarrow B$ y $f : B \longrightarrow C$, tales que $f \circ g$ es inyectiva, entonces g es inyectiva. Esto se sigue ya que si $g(x_1) = g(x_2)$, necesariamente $fg(x_1) = fg(x_2)$ y $x_1 = x_2$.

También si $f : A \longrightarrow B$ y $g : B \longrightarrow C$ son tales que $g \circ f$ es suprayectiva, entonces g es suprayectiva. Esto se sigue ya que dado $z \in C$, $\exists\, x \in A \ni g \circ f(x) = z$, por lo cual $g(y) = z$, donde $f(x) = y$.

Por otra parte, si $f : A \longrightarrow B$ es suprayectiva y $g, h : B \longrightarrow C$ son tales que $g \circ f = h \circ f$, entonces $g = h$. Para probar esto se toma $y \in B$, entonces $\exists\, x \in A$ tal que $f(x) = y$ y se tiene $gf(x) = hf(x)$, *i.e.*, $g(y) = h(y)$.

EJERCICIO 2.6

1. Pruebe que si $f : B \longrightarrow C$ es inyectiva y $g, h : A \longrightarrow B$ son tales que $f \circ g = f \circ h$, entonces $g = h$.

2.7. Ordenaciones con repetición (versión formal)

Definición 27 *Sea A un conjunto de n elementos, las ordenaciones con repetición de A tomadas de m en m son las funciones $f : I_m \longrightarrow A$, donde $I_m = \{1, 2, ..., m\}$.*

Recordando las señales telegráficas corta y larga: $\cdot -$, las de tres sonidos pueden interpretarse en términos de funciones, como sigue

$$\begin{pmatrix} 1 & 2 & 3 \\ \cdot & \cdot & \cdot \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ \cdot & \cdot & - \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ \cdot & - & \cdot \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ \cdot & - & - \end{pmatrix}$$
$$\begin{pmatrix} 1 & 2 & 3 \\ - & \cdot & \cdot \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ - & \cdot & - \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ - & - & \cdot \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ - & - & - \end{pmatrix}.$$

Como un segundo ejemplo consideremos las ordenaciones con repetición del conjunto $A = \{x, y, z\}$ tomadas de 2 en 2, éstas son:

$$\begin{pmatrix} 1 & 2 \\ x & x \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ x & y \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ x & z \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ y & x \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ y & y \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ y & z \end{pmatrix}$$
$$\begin{pmatrix} 1 & 2 \\ z & x \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ z & y \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ z & z \end{pmatrix}.$$

Obsérvese que estas parejas son, en cierta manera, los elementos del producto cartesiano $A \times A$. Análogamente, las ordenaciones con repetición tomadas de 3 en 3 del conjunto A son, en cierta manera, los elementos de $A \times A \times A$.

Algunas veces en lugar de escribir

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ a_1 & a_2 & a_3 & \cdots & a_n \end{pmatrix}$$

se escribe simplemente $(a_1, a_2, a_3, ..., a_n)$.

Definición 28 *Sea $f : A \longrightarrow B$, $g : C \longrightarrow B$ funciones, tales que $A \subset C$. Supóngase también que*

$$f(x) = g(x) \quad \forall\, x \in A.$$

Bajo estas hipótesis se dice que g es una extensión de f y que g restringida a A, denotado por $g \,|\, A$, es igual a f (se escribe $g \,|\, A = f$).

Por ejemplo, si $f : \mathbb{Z} \longrightarrow \mathbb{Z}$, donde $f(n) = n^2$ y $g : \mathbb{R} \longrightarrow \mathbb{R}$, $g(x) = x^2$, entonces $g \,|\, \mathbb{Z} = f$.

Consideramos ahora un ejemplo que nos servirá para entender la prueba de que $OR_n^m = n^m$.

Sean A un conjunto de 10 elementos, a y b dos elementos de A no necesariamente distintos y $f : I_2 \longrightarrow A \ni f(1) = a$ y $f(2) = b$. Es claro que el número de funciones que extienden f a funciones con dominio I_3 son 10, ya que hay 10 maneras de elegir la imagen de 3.

Teorema 2.7.1 *Sea A un conjunto de n elementos, y $m \in \mathbb{N}$, entonces*

$$OR_n^m = n^m.$$

Demostración. Inducción sobre m, donde n es fija.

Si $m = 1$

$$OR_n^1 = n,$$

ya que es el número de funciones de I_1 en A.

Suponemos cierto el teorema para $m - 1$, se afirma que

$$OR_n^m = n\, OR_n^{m-1}.$$

La afirmación implica el resultado ya que $n \cdot n^{m-1} = n^m$. Para probar la afirmación basta probar que dada $f : I_{m-1} \longrightarrow A$ existen n funciones que extienden f a funciones de I_m en A, esto es claro ya que la imagen de m puede ser cualquier elemento de A. □

Definición 29 *Sea A un conjunto, el producto cartesiano*

$$\underbrace{A \times A \times \cdots \times A}_{n-veces},$$

denotado por A^n, es el conjunto cuyos elementos son las n-adas de elementos de A, es decir, elementos de la forma

$$(a_1, a_2, ..., a_n), \quad a_i \in A, \quad i \in \{1, 2, ..., n\}.$$

Por ejemplo: si $n = 2$, son las parejas ordenadas. Si $n = 3$ son ternas ordenadas, etcétera. Si $A = \mathbb{R}$, $\mathbb{R}^3 = \{(x, y, z) \,|\, x, y, z \in \mathbb{R}\}$.

Definición 30 *Una correspondencia biunívoca entre 2 conjuntos A y B es una función biyectiva $f : A \longrightarrow B$.*

Las observaciones anteriores muestran que si A es un conjunto de n elementos, la cardinalidad de A^m es precisamente el número OR_n^m. Esto se sigue ya que hay una correspondencia biunívoca entre las ordenaciones con repetición de A tomadas de m en m con los elementos de A^m.

Si $A = \{a_1, ..., a_n\}$, esta correspondencia está dada por

$$\begin{pmatrix} 1 & 2 & \cdots & m \\ a_{i_1} & a_{i_2} & \cdots & a_{i_m} \end{pmatrix} \longleftrightarrow \left(a_{i_1}, a_{i_2}, ..., a_{i_m} \right),$$

$a_{i_j} \in A \qquad \forall\, j.$

2.8. Ordenaciones (versión formal)

Definición 31 *Dado A un conjunto de n elementos y $m \leq n$, las ordenaciones de los elementos de A tomadas de m en m, son las funciones inyectivas del conjunto I_m en A.*

Observemos que la inyectividad es la condición que establece que no haya repeticiones. Como ejemplo, consideremos las ordenaciones del conjunto $A = \{a, b, c, d, e\}$, tomadas de tres en tres. Éstas se pueden pensar también como palabras con tres letras distintas.

Un método para encontrarlas es considerar, por ejemplo, la función inyectiva $f : I_2 \longrightarrow A$ dado por

$$\begin{pmatrix} 1 & 2 \\ a & e \end{pmatrix}.$$

Ahora la pregunta relevante es ¿cuántas funciones inyectivas hay de I_3 en A que extiendan a f? Éstas son

$$\begin{pmatrix} 1 & 2 & 3 \\ a & e & x \end{pmatrix},$$

donde $x = b, c, d,$ es decir, hay 5-2 funciones. Se concluye que por cada función inyectiva de I_2 en A hay 3 de I_3 en A.

Teorema 2.8.1 *Sean* $n, m \in \mathbb{N},\ m \leq n,$ *entonces*

$$O_n^m = \frac{n!}{(n-m)!}.$$

Demostración. Se usará inducción sobre m. Sea A un conjunto de n elementos. En el caso $m = 1$, es evidente que hay n funciones inyectivas de I_1 en A y se sigue que

$$O_n^1 = \frac{n!}{(n-1)!}.$$

Suponemos cierta la fórmula para $m-1$

$$O_n^{m-1} = \frac{n!}{(n-(m-1))!}.$$

El ejemplo anterior ilustra el paso inductivo. Dada cualquier función inyectiva f de I_{m-1} en A, f se puede extender a $n-(m-1)$ funciones inyectivas de I_m en A, ya que al elemento m se le puede asociar cualquier elemento que no sea la imagen de $1, 2, ..., m-1$ (por la inyectividad).

En consecuencia

$$O_n^m = (n-(m-1))\, O_n^{m-1} = (n-(m-1))\frac{n!}{(n-(m-1))!}$$

$$= (n-m+1)\frac{n!}{(n-m+1)!} = \frac{n!}{(n-m)!}$$

□

La fórmula en el Teorema 2.7.1 resuelve el siguiente problema: ¿cuántas placas de automóvil hay que consten de 3 letras y 2 cifras?

Considerando 27 letras en el alfabeto, placas de 3 letras hay $OR_{27}^3 = 27^3$, y de 2 cifras hay $OR_{10}^2 = 10^2$. Por cada ordenación (con repetición) en la primera lista, esto es, la de las letras, se puede generar OR_{10}^2 placas distintas, lo cual muestra el carácter multiplicativo y es claro que el número total de placas es

$$OR_{27}^3\ OR_{10}^2 = 27^3 \cdot 10^2$$

Ejemplos

1. ¿Cuántos números telefónicos de 6 cifras hay que comiencen con $5, 7, 2, 6$ u 8?

   De 5 cifras hay $OR_{10}^5 = 10^5$, considerando la 1a cifra se tienen $5 \cdot 10^5$.

2. ¿Cuántas placas de automóvil hay que consten de dos letras y tres cifras, si la primera letra es K y la segunda una letra de la A a la D?

   4×10^3.

3. ¿Cuántas placas de siete cifras distintas pueden formarse si la 1a la 2a y la 5a son cifras pares?

   Las posibilidades de ordenaciones tomadas de 3 en 3 con cifras pares son O_5^3 (lugares 1, 2 y 5). Fijando cualquiera de ellas, las posibilidades de ordenaciones que existen son O_7^4 (puesto que tres números ya aparecen en los lugares 1, 2 y 5) $\therefore$ el número total es $O_5^3 \ O_7^4$.

EJERCICIOS 2.8

1. Sean $B = \{x, y, z\}$ y $f : I_3 \longrightarrow B$ tal que manda 1 y 2 en x y 3 en y, ¿cuántas extensiones hay de f a I_5?
2. ¿Cuántos números telefónicos hay con tres números, no necesariamente distintos, si exactamente dos de ellos son impares?
3. ¿Cuántas placas de automóvil hay de tres cifras y cuatro letras, si las cifras son distintas?
4. ¿Cuántas cifras de cuatro dígitos distintos hay, si exactamente dos de ellos son impares?
5. ¿Cuántas placas de cuatro números hay, que tengan al menos dos números iguales?

2.9. Permutaciones (versión formal)

Definición 32 *Las permutaciones de un conjunto A son las funciones biyectivas de A en A.*

Se demostró que si A es finito y $f : A \longrightarrow A$ es inyectiva, entonces f es biyectiva.

Sean $A = \{a_1, a_2, .., a_n\}$ y $f : A \longrightarrow A$ una permutación, entonces se puede identificar a f con la ordenación $g : I_n \longrightarrow A$ dada por

$$g = \begin{pmatrix} 1 & 2 & \cdots & n \\ f(a_1) & f(a_2) & \cdots & f(a_n) \end{pmatrix},$$

y viceversa dada cualquier función inyectiva de I_n en A se le puede asociar una única permutación de A.

Por ejemplo, si $A = \{a, b, c\}$, las seis permutaciones de A son:

$$\begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix} \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix} \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix} \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix} \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix} \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}$$

que se pueden indentificar con las ordenaciones

$$\begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ a & c & b \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ b & a & c \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ b & c & a \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ c & a & b \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ c & b & a \end{pmatrix}.$$

Dada esta correspondencia biunívoca se sigue del Teorema 2.8.1 que

$$P_n = O_n^n = \frac{n!}{(n-n)!} = n!.$$

2.10. Combinaciones y coeficientes binomiales

Recordamos la definición de combinación.

Definición 33 *Sea A un conjunto de n elementos y $0 \leq m \leq n$. A los subconjuntos de A que tienen m elementos se les llama combinaciones de A tomadas de m en m, y el número de éstas se denota por C_n^m.*

Probaremos ahora que C_n^m es precisamente el coeficiente binomial $\frac{n!}{(n-m)!m!}$, lo cual justifica el uso doble de este símbolo.

Teorema 2.10.1 *Sea A un conjunto de n elementos, entonces*

$$C_n^m P_m = O_n^m.$$

Demostración. Sea S el conjunto de las ordenaciones de los elementos de A tomadas de m en m y T el conjunto de las combinaciones de los elementos de A tomadas de m en m. Obsérvese que $\#S = O_n^m$ y $\#T = C_n^m$.

Sea $\psi : S \longrightarrow T$ dada por

$$\psi\left[\begin{pmatrix} 1 & 2 & \cdots & m \\ a_1 & a_2 & \cdots & a_m \end{pmatrix}\right] = \{a_1, a_2, ..., a_m\},$$

donde $a_j \in A \quad \forall j$ y $a_i \neq a_j$ si $i \neq j$, esencialmente la función ψ 'olvida' el orden. Claramente ψ es suprayectiva, ya que dados m elementos distintos

en A se puede construir una función inyectiva de I_m en A cuyas imágenes sean estos elementos.

Además existen exactamente P_m ordenaciones a las que se les asocia la misma combinación. En consecuencia

$$C_n^m P_m = O_n^m.$$

□

Obsérvese que el teorema anterior establece que

$$C_n^m = \frac{O_n^m}{P_m} = \frac{n!}{(n-m)!m!}.$$

Un ejemplo que ilustra la última parte de la prueba del teorema anterior es el siguiente: sea $A = \{a, b, c, ..., z\}$ y consideremos las ordenaciones asociadas a la combinación $\{a, b, c\}$, éstas son:

$$\begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ a & c & b \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ b & a & c \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ b & c & a \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ c & a & b \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ c & b & a \end{pmatrix}.$$

A los números C_n^m se les llama coeficientes binomiales ya que aparecen en el desarrollo del binomio de Newton

$$(a+b)^n = C_n^o a^n + C_n^1 a^{n-1} b^1 + \cdots + C_n^n b^n.$$

Obsérvese que $C_n^o = 1$, ya que el número de subconjuntos que no tienen elementos del conjunto $A = \{a_1, a_2, ..., a_n\}$ es uno, el vacío.

Teorema 2.10.2 *Sean m, n enteros no negativos, tales que $m \leq n$, entonces*

$$C_n^m = C_n^{n-m}.$$

Demostración. Sea A un conjunto de n elementos, S las combinaciones de A tomadas de m en m y T las combinaciones de A tomadas de $n-m$ en $n-m$.

Definimos una función $\psi : S \longrightarrow T$ de la siguiente manera: a cada combinación $\{a_1, a_2, ..., a_m\}$ en S se le asocia $A - \{a_1, a_2, ..., a_m\}$ en T, *i.e.*, su complemento.

Esta función ψ es claramente inyectiva, ya que dos subconjuntos distintos de A tienen complementos distintos. También es suprayectiva, ya que si $B \subset A$, entonces $\psi(B^c) = B$.

Por lo tanto, ψ es biyectiva y se sigue el teorema

$$\#S = \#T, \quad i.e., \quad C_n^m = C_n^{n-m}.$$

□

Otra prueba más simple del teorema anterior, pero menos ilustrativa, es la siguiente

$$\frac{n!}{(n-m)!\,m!} = \frac{n!}{(n-(n-m))!\,(n-m)!}.$$

Definición 34 *Sea B un subconjunto de A. Se define la siguiente función $f_B : A \longrightarrow \{0,1\}$, llamada característica, como sigue*

$$f_B(x) = \begin{cases} 1 & \text{si } x \in B, \\ 0 & \text{si } x \notin B. \end{cases}$$

Obsérvese que la asociación

$$B \longrightarrow f_B$$

es una biyección entre los subconjuntos de A y las funciones características definidas en A, *i.e.*, las funciones de A en $\{0,1\}$. Si $B_1 \neq B_2$, es claro que $f_{B_1} \neq f_{B_2}$. Más aún, dada una función característica $g : A \longrightarrow \{0,1\}$, ésta viene del subconjunto que la define, *i.e.*, si $g^{-1}(1) = B$, se sigue que $g = f_B$.

Teorema 2.10.3 *Sea n un número natural, entonces*

$$C_n^0 + C_n^1 + \cdots + C_n^n = 2^n.$$

Demostración. Obsérvese que el miembro izquierdo es el número total de subconjuntos de un conjunto A de n elementos, por lo que se sigue de la observación anterior que es el número de funciones de A en $\{0,1\}$. Se afirma que este número es 2^n, esto se puede probar por inducción sobre n. Si A tiene un elemento hay 2 funciones de A en $\{0,1\}$.

Suponiendo cierta la afirmación para $n-1$, *i.e.*, si $A = \{a_1, a_2, ..., a_{n-1}\}$ hay 2^{n-1} funciones de A en $\{0,1\}$. Se sigue entonces que cada una de estas funciones tiene dos extensiones a funciones de $\{a_1, a_2, ..., a_n\}$ en $\{0,1\}$, al elemento a_n se le puede asignar el 0, o el 1.

□

Otra demostración más breve se sigue del teorema del binomio

$$(1+1)^n = \sum_{j=0}^{n} C_n^j.$$

Terminamos esta sección con una segunda demostración de la fórmula del triángulo de Pascal

$$C_n^{r-1} + C_n^r = C_{n+1}^r.$$

Sea $A = \{1, 2, ..., n+1\}$. Los subconjuntos de A con r elementos (que son un total de C_{n+1}^r) son de 2 tipos:

a) los que no contienen al elemento $n+1$, de estos hay C_n^r,

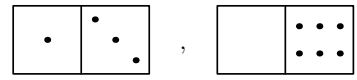
b) los que contienen al elemento $n+1$, éstos quedan determinados por los subconjuntos de $\{1, 2, ..., n\}$ que contienen $r-1$ elementos, de estos hay C_n^{r-1}.

En consecuencia, se sigue la fórmula de Pascal.

1	C_0^0
1 1	C_1^0 C_1^1
1 2 1	C_2^0 C_2^1 C_2^2
1 3 3 1	C_3^0 C_3^1 C_3^2 C_3^3
1 4 6 4 1	C_4^0 C_4^1 C_4^2 C_4^3 C_4^4
1 5 10 10 5 1	

Figura 2.8: Triángulo de Pascal

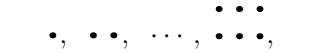
El dominó consta de 28 fichas cada una con dos lados, cada lado es blanco, o tiene $1, 2, 3, 4, 5$ o 6 puntos, por ejemplo,



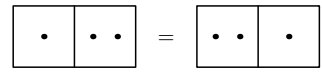
1 3 0 6

Figura 2.9: Ejemplo de fichas de dominó

Veamos por qué son 28 fichas. En efecto, una ficha se puede pensar como una ordenación con repetición de 2 lugares con 7 símbolos: blanco, o



i.e., son 7^2, sin embargo las que no son dobles se repiten, por ejemplo,



Por lo que el número de fichas es

$$\frac{49-7}{2}+7=21+7=28.$$

Ejemplos

1. De un grupo de 20 personas debe elegirse un comité de 5 miembros en el que debe estar Juan o Pedro, pero no ambos. ¿Cuántos comités pueden ser electos?

   Excluyendo a Juan y a Pedro, hay C_{18}^4 maneras de elegir los otros 4 miembros, *i.e.*, hay $2 \cdot C_{18}^4$ maneras.

2. Consideramos la baraja simplificada, es decir, con solamente 3 números y 2 palos. ¿Cuántas manos de 3 cartas hay que no tengan 2 cartas del mismo número?

   La respuesta se obtiene al restar del número total de manos, las manos que tienen un par. Ahora, hay C_6^3 manos posibles. Por otro lado, las manos con un par se obtienen al elegir un número entre 3, *i.e.*, hay C_3^1 maneras, la 3a carta se puede elegir de 4 maneras distintas. Por lo que el número total es

   $$C_6^3 - 3 \cdot 4 = \frac{6!}{3!3!} - 3 \cdot 4 = 5 \cdot 4 - 3 \cdot 4 = 8.$$

3. El dominó consta de 28 fichas y una mano de dominó consta de 7 fichas ¿Cuántas manos de dominó hay?

   El número de manos son C_{28}^7.

4. Sea A un conjunto con n elementos, $n \geq 5$ y x_0 un elemento de A. ¿Cuántos subconjuntos de A hay con 5 elementos, tales que contengan al elemento x_0?

   El número es C_{n-1}^4, ya que se toman de un conjunto de $n-1$ elementos, 4 de ellos.

5. ¿Cuántas manos de poker hay que no tengan dos cartas del mismo número?

   Se seleccionan primero los 5 números distintos, éstos son tantos como C_{13}^5. Ahora dados 5 números distintos, las posibilidades para los palos son las ordenaciones con repetición de un conjunto de 4 elementos $\{\blacklozenge, \heartsuit, \clubsuit, \spadesuit\}$ tomados de 5 en 5. Fijando, por ejemplo 7,8,9,Q,K las posibles manos con estos números corresponden a las palos distintos, funciones de I_5 en $\{\blacklozenge, \heartsuit, \clubsuit, \spadesuit\}$. Por lo tanto hay $C_{13}^5 \cdot OR_4^5 = \frac{13!}{5!8!} \cdot 4^5$.

6. ¿Cuántas manos de poker hay que tengan exactamente un par?

   Elegir un par es elegir un número de 13, *i.e.*, C_{13}^1, ahora hay 4 palos y se deben tomar 2 de 4, *i.e.,* C_4^2. Fijando este par se piensa cuántas formas hay de elegir las otras 3 cartas, para esto hay C_{12}^3 maneras de elegir 3 números distintos (13 - 1 = 12), y para los palos, siguiendo la técnica del ejemplo 5, hay OR_4^3 maneras de elegir los palos. Por lo tanto el número total es

$$C_{13}^1 C_4^2 C_{12}^3 OR_4^3 = 13 \cdot \frac{4!}{2!2!} \cdot \frac{12!}{9!3!} \cdot 4^3 = 13 \cdot 12 \cdot 11 \cdot 10 \cdot 4^3.$$

7. ¿Cuántas manos de poker hay que tengan 2 pares distintos, que no sea full (es decir una tercia y un par)?

   Hay C_{13}^2 maneras de elegir 2 números distintos de 13, como en el ejemplo anterior hay C_4^2 maneras de elegir 2 palos para un par y C_4^2 para el otro. Finalmente hay 44 = 52 - 8 maneras de elegir la quinta carta, por lo que el número de manos es

$$44(C_4^2)^2 C_{13}^2 = 44 \cdot \frac{4!}{2!2!} \cdot \frac{4!}{2!2!} \cdot \frac{13!}{11!2!} = 22 \cdot 36 \cdot 13 \cdot 12.$$

8. ¿Cuántas manos de poker hay que tengan al menos 3 cartas del mismo número?

   Calculamos independientemente el número de tercias y de poker (4 cartas iguales). Hay 13 maneras de elegir un número, $C_4^3 = C_4^1$ maneras de elegir 3 palos de 4, y C_{48}^2 maneras de elegir las 2 cartas restantes (de 2 números distintos a la tercia), *i.e.*, el número de tercias es

$$13 \cdot 4 \cdot \frac{48!}{46!2!} = 13 \cdot 2 \cdot 48 \cdot 47.$$

El número de manos de poker es $13 \cdot 48$ por lo que el número total es

$$13 \cdot 2 \cdot 48 \cdot 47 + 13 \cdot 48 = 13 \cdot 48 \cdot 95.$$

9. ¿Cuántas diagonales se pueden trazar en un polígono regular de n lados?

   Por cada dos vértices se puede trazar una diagonal, sin embargo el segmento de recta que une un vértice con su adyacente no es diagonal

   $\therefore$ el número es

$$C_n^2 - n = \frac{n!}{(n-2)!2!} - n = \frac{n(n-1)}{2} - n = n\left(\frac{n}{2} - \frac{3}{2}\right) = \frac{n(n-3)}{2}.$$

10. ¿Cuántas manos de dominó hay que tengan exactamente 4 fichas dobles?

    C_7^4 son las posibilidades de elegir 4 números de 7 (fichas dobles), habiendo elegido las 4 dobles las 3 restantes se pueden elegir entre 21 (las otras 3 dobles se excluyen).

    $\therefore$ la respuesta es $C_7^4\, C_{21}^3$.

11. ¿Cuántas manos de dominó hay que tengan por lo menos 3 fichas dobles?

    Exactamente 3 dobles $C_7^3\, C_{21}^4$, exactamente 4 hay $C_7^4\, C_{21}^3$, etcétera, por lo que el total es

$$C_7^3\, C_{21}^4 + C_7^4\, C_{21}^3 + C_7^5\, C_{21}^2 + C_7^6\, C_{21}^1 + 1.$$

EJERCICIOS 2.10

1. ¿De cuántas maneras se pueden distribuir 4 libros distintos entre 2 estudiantes?
2. ¿Cuántas manos de poker hay que tengan full, *i.e.*, tercia y par?
3. ¿Cuántas manos de poker hay que tengan flor (todos los números del mismo palo)?
4. ¿Cuántas manos de poker hay que tengan flor imperial, *i.e.*, todas las cartas del mismo palo, y los números cíclicamente consecutivos?
5. ¿Cuántos números telefónicos con cuatro cifras, no necesariamente distintas hay, si exactamente dos de ellas son impares?

Capítulo 3

El espacio vectorial $\mathbb{R}^n$

3.1. Vectores y sus operaciones

Definición 35 *El producto cartesiano* $\mathbb{R} \times \mathbb{R}$ *es el conjunto de parejas ordenadas* (a, b), *donde* $a, b \in \mathbb{R}$, *se le designa como* $\mathbb{R}^2$ *y se le identifica con los puntos del plano.*

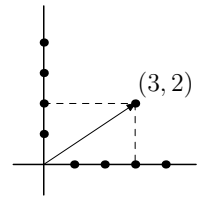


Figura 3.1: Puntos del plano cartesiano

A los elementos del plano o puntos de $\mathbb{R}^2$ se les llama *vectores* y a los números reales se les llama *escalares*. Por ejemplo, el vector $(3, 2)$ es el punto descrito en la Figura 3.1. Se definen las siguientes operaciones en $\mathbb{R}^2$.

1) SUMA DE VECTORES

$$(a, b) + (c, d) = (a + c,\, b + d) \quad \forall\, (a, b),\, (c, d) \in \mathbb{R}^2.$$

2) PRODUCTO DE UN ESCALAR POR UN VECTOR

$$\lambda(a, b) = (\lambda\, a, \lambda\, b) \quad \forall\, \lambda \in \mathbb{R} \quad \text{y} \quad \forall\, (a, b) \in \mathbb{R}^2.$$

Si (a, b) es un vector, a la primera coordenada a se le llama abcisa y a la segunda coordenada b se la llama ordenada.

Ejemplos

$(-1, 2) + (1, 2) = (0, 4)$,

$(a, 0) + (0, b) = (a, b)$,

$3(1, 2) = (3, 6)$,

$b(0, 1) = (0, b)$.

Obsérvese que $(a, b) = a(1, 0) + b(0, 1)$. Exhibimos ahora una interpretación geométrica de la adición, para esto recordamos dos hechos de la geometría elemental.

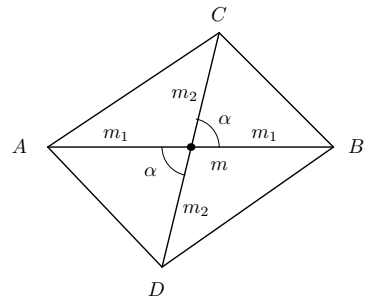


Figura 3.2: Un cuadrilátero cuyas diagonales se encuentran en su punto medio es un paralelogramo

I. *Si las diagonales de un cuadrilátero se intersecan en su punto medio, entonces es un paralelogramo.*

Para probar este hecho, nótese primero que el triángulo $\triangle A\,m\,C$ es igual al triángulo $\triangle D\,m\,B$ y también que $\triangle A\,m\,D = \triangle C\,m\,B$. Esto se sigue de que tienen un ángulo común delimitado por lados iguales. Por consiguiente, los ángulos δ, que aparecen en la Figura 3.3, son iguales, y esto implica que $\overline{AC}$ es paralela a $\overline{BD}$, y que $\overline{AD}$ también lo es a $\overline{CB}$. De otra manera, por ejemplo, se toma la paralela a $\overline{AC}$ que pasa por D, y al considerar ángulos alternos internos se llega a una contradicción.

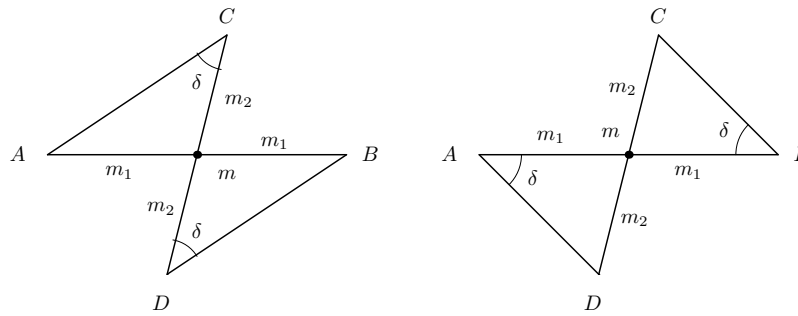


Figura 3.3: Prueba de la afirmación I

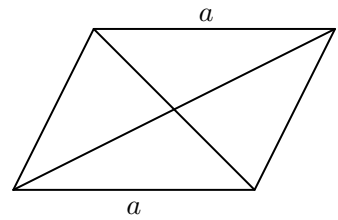


Figura 3.4: En un paralelogramo las diagonales se intersectan en el punto medio

Obsérvese que también es cierto el recíproco, es decir, en un paralelogramo las diagonales se intersecan en su punto medio. A esto se llega por el hecho de que los 4 triángulos que se generan son iguales, dos a dos, al ser semejantes y tener un lado común, véase la Figura 3.4.

II. *Si* $A_1 = (x_1, y_1)$ *y* $A_2 = (x_2, y_2)$ *son 2 puntos del plano, entonces el punto medio del segmento* $\overline{A_1A_2}$ *está dado por*

$$m = \left(\frac{x_1 + x_2}{2}, \ \frac{y_1 + y_2}{2} \right).$$

Por semejanza, en los tríangulos descritos en la Figura 3.5 se sigue que

$$\frac{y_2 - y_1}{y_2 - m_2} = 2, \quad i.e., \quad y_2 - y_1 = 2\,(y_2 - m_2),$$

donde $m = (m_1, m_2)$. En consecuencia, $2\,m_2 = y_1 + y_2$. Un argumento similar prueba que también $2\,m_1 = x_1 + x_2$.

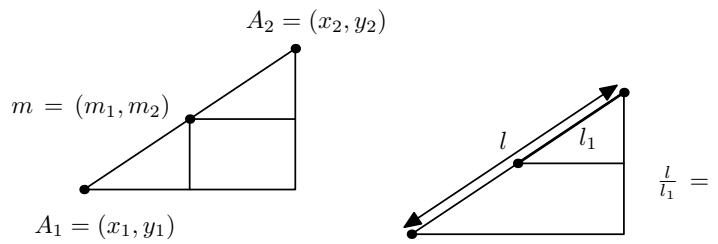


Figura 3.5: Punto medio de un segmento

Interpretación geométrica de la adición de vectores

Dados los vectores $P = (a, b)$, $Q = (c, d)$, *el cuadrilátero formado por los vectores* O, P, Q *y* $P + Q$ *es un paralelogramo.*

Esto sucede, ya que si

$$P + Q = R = (a + c, b + d),$$

los puntos medios de $\overline{OR}$ y de $\overline{PQ}$ son iguales, ya que éstos están dados por

$$\left(\frac{a + c + 0}{2}, \frac{b + d + 0}{2}\right) \qquad \text{y} \qquad \left(\frac{a + c}{2}, \frac{b + d}{2}\right),$$

respectivamente, véase la Figura 3.6. Obsérvese que la suma de los vectores, $P + Q$ es el vector determinado por la diagonal desde el origen a $P + Q$, por lo que se le llama la resultante.

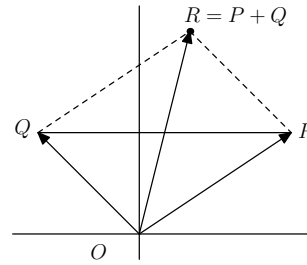


Figura 3.6: Interpretación geométrica de la adición

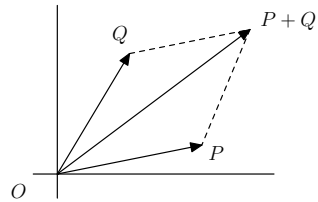


Figura 3.7: La resultante

INTERPRETACIÓN GEOMÉTRICA DEL PRODUCTO DE UN ESCALAR POR UN VECTOR

Probamos primero un resultado. Recordamos que dados 2 puntos $P = (x_1, y_1)$ y $Q = (x_2, y_2)$, se define su distancia como

$$d(P, Q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

La motivación para esta definición es, por supuesto, el teorema de Pitágoras.

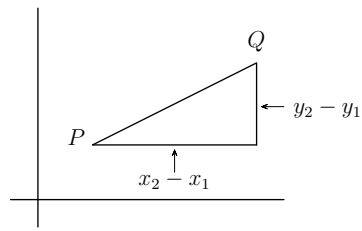


Figura 3.8: Distancia entre dos puntos

Proposición 3.1.1 *Sean P_1, P_2, P_3 3 puntos del plano tales que $P_2 \in \overline{P_1P_3}$, entonces*

$$d(P_1, P_2) + d(P_2, P_3) = d(P_1, P_3).$$

Demostración. Sean $P_2 - P_1 = (x_1, y_1)$ y $P_3 - P_2 = (x_2, y_2)$, entonces

$$d(P_1, P_2) = \sqrt{{x_1}^2 + {y_1}^2} \quad \text{y} \quad d(P_2, P_3) = \sqrt{{x_2}^2 + {y_2}^2}.$$

En virtud de la semejanza de triángulos en la Figura 3.9, se sigue que

$$\frac{y_1}{x_1} = \frac{y_2}{x_2} = \frac{y_1 + y_2}{x_1 + x_2} = \lambda.$$

En consecuencia, se llega al resultado, *i.e.*,

$$\sqrt{{x_1}^2 + {y_1}^2} + \sqrt{{x_2}^2 + {y_2}^2} = \sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2},$$

ya que éste es equivalente a la ecuación

$$x_1\sqrt{1 + \lambda^2} + x_2\sqrt{1 + \lambda^2} = (x_1 + x_2)\sqrt{1 + \lambda^2}.$$

□

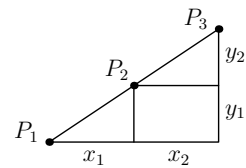


Figura 3.9: Proporcionalidad de las distancias en tres puntos alineados

El recíproco también es cierto como se muestra en el siguiente resultado.

Proposición 3.1.2 *Si se cumple*

$$d(P_1, P_2) + d(P_2, P_3) = d(P_1, P_3), \tag{3.1}$$

entonces $P_2 \in \overline{P_1P_3}$.

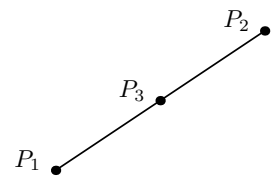


Figura 3.10: Prueba de la Proposición 3.1.2

Demostración. Si $P_2 \notin \overline{P_1P_3}$ probamos que no se cumple la igualdad (3.1)

Si P_2 está en la misma línea que P_1 y P_3 y $P_2 \notin \overline{P_1P_3}$, por ejemplo, $P_3 \in \overline{P_1P_2}$. No se puede cumplir (3.1), ya que en virtud de la Proposición 3.1.1 se sigue que $d(P_1, P_3) < d(P_1, P_2)$, véase la Figura 3.10. El otro caso es análogo.

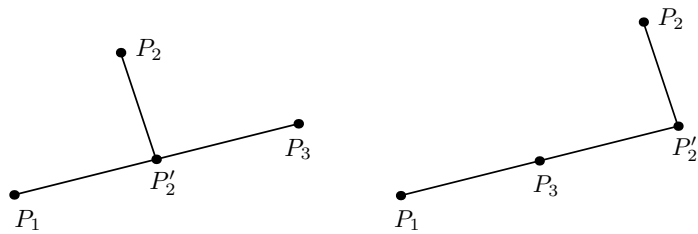


Figura 3.11: Prueba de la Proposición 3.1.2

Si P_2 no está en la recta determinada por el segmento $\overline{P_1P_3}$, se tienen otros dos casos, como se muestra en la Figura 3.11.

En el primer caso de la figura se cumple la igualdad para P_2', y por lo tanto no se cumple para P_2 (por el teorema de Pitágoras). En el segundo caso de la figura tampoco, ya que $d(P_1, P_2) > d(P_1, P_2') > d(P_1, P_3)$. Un argumento similar muestra también que no se cumple la igualdad, si P_2 se proyecta del lado de P_1. □

Proposición 3.1.3 *(Interpretación geométrica del producto de un escalar por un vector). Sean $P \neq O$ y $Q = \lambda P$, puntos del plano, donde $O = (0,0)$ y $\lambda \in \mathbb{R}$, entonces*

a) si $\lambda \geq 1$, P *está en el segmento* $\overline{OQ}$,

b) si $0 \leq \lambda \leq 1$, $Q \in \overline{OP}$,

c) si $\lambda \leq 0$, $O \in \overline{PQ}$.

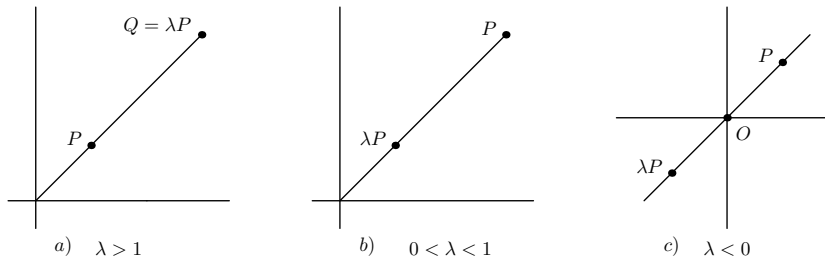


Figura 3.12: 3 casos de la Proposición 3.1.3

DEMOSTRACIÓN. Sean $P = (a, b)$ y $|P| = d(O, P)$, *i.e.*, $|P| = \sqrt{a^2 + b^2}$, entonces

$$d(O, \lambda P) = \sqrt{(\lambda a)^2 + (\lambda b)^2} = |\lambda||P|,$$

y también

$$d(P, Q) = \sqrt{(a - \lambda a)^2 + (b - \lambda b)^2} = |1 - \lambda||P|.$$

Consideramos casos.

Caso 1: $\lambda \geq 1$. Se tiene

$$d(O, P) + d(P, Q) = |P| + (\lambda - 1)|P| = \lambda|P| = |\lambda||P| = d(O, Q),$$

ya que $\lambda \geq 1$, y por lo tanto $P \in \overline{OQ}$ (por la Proposición 3.1.2).

Caso 2: $0 \leq \lambda \leq 1$. Se tiene

$$d(O, Q) + d(Q, P) = \lambda|P| + (1 - \lambda)|P| = |P| = d(O, P)$$

y $Q \in \overline{OP}$.

Caso 3: $\lambda \leq 0$. Se tiene

$$d(Q, O) + d(O, P) = -\lambda|P| + |P| = (1 - \lambda)|P| = |1 - \lambda||P| = d(P, Q)$$

$\therefore \quad O \in \overline{QP}$.

□

3.2. $\mathbb{R}^n$

Definición 36 *El espacio vectorial* $\mathbb{R}^n$ *es el conjunto de todas las colecciones ordenadas de n números reales* $(a_1, a_2, ..., a_n)$, *i.e.,*

$$\mathbb{R}^n = \{(a_1, a_2, ..., a_n) \,|\, a_i \in \mathbb{R}\ \forall i \in \{1, 2, ..., n\}\}.$$

Se tienen 2 operaciones:

SUMA

$$(a_1, a_2, ..., a_n) + (b_1, b_2, ..., b_n) = (a_1 + b_1, a_2 + b_2, ..., a_n + b_n).$$

PRODUCTO POR ESCALARES

$$\lambda(a_1, a_2, ..., a_n) = (\lambda a_1, \lambda a_2, ..., \lambda a_n).$$

A los elementos $(a_1, a_2, ..., a_n)$ se les llama vectores, a los números reales se les llama escalares.

PROPIEDADES

1. Conmutatividad

$$A + B = B + A \qquad \forall\, A, B \in \mathbb{R}^n.$$

2. Asociatividad

$$(A + B) + C = A + (B + C) \qquad \forall A, B, C \in \mathbb{R}^n.$$

3. Existencia del neutro aditivo

$$A + \overline{0} = A \qquad \forall A \in \mathbb{R}^n \quad \text{donde} \quad \overline{0} = (0, 0, ..., 0).$$

4. Existencia del inverso aditivo

$\forall A \in \mathbb{R}^n$ existe $B \in \mathbb{R}^n$ ∋

$$A + B = \overline{0}, \quad \text{si} \quad A = (a_1, a_2, ..., a_n),\ B = (-a_1, -a_2, ..., -a_n),$$

el inverso aditivo B se denota por $-A$.

Estas cuatro propiedades se derivan de las mismas leyes para los números reales, por ejemplo,

$$\begin{aligned}(a_1, a_2, a_3) + (b_1, b_2, b_3) &= (a_1 + b_1, a_2 + b_2, a_3 + b_3)\\ &= (b_1 + a_1, b_2 + a_2, b_3 + a_3)\\ &= (b_1, b_2, b_3) + (a_1, a_2, a_3).\end{aligned}$$

5. Si $A \in \mathbb{R}^n$ y $\lambda, \mu \in \mathbb{R}$, entonces $\lambda(\mu A) = (\lambda\mu)A$.

Demostración. Sea $A = (a_1, a_2, ..., a_n)$, entonces

$$\lambda(\mu A) = \lambda(\mu a_1, ..., \mu a_n) = (\lambda\mu a_1, ..., \lambda\mu a_n) = (\lambda\mu)A.$$

6. Si $A, B \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$, entonces $\lambda(A + B) = \lambda A + \lambda B$.

Demostración. Sean $A = (a_1, a_2, ..., a_n)$ y $B = (b_1, b_2, ..., b_n)$

$$\begin{aligned}\lambda(A + B) &= \lambda[(a_1, ..., a_n) + (b_1, ..., b_n)]\\ &= (\lambda(a_1 + b_1), ..., \lambda(a_n + b_n))\\ &= (\lambda a_1 + \lambda b_1, ..., \lambda a_n + \lambda b_n)\\ &= \lambda A + \lambda B.\end{aligned}$$

7. Si $A \in \mathbb{R}^n$, y $\lambda, \mu \in \mathbb{R}$, entonces $(\lambda + \mu)A = \lambda A + \mu A$.

   Demostración.

$$\begin{aligned}(\lambda+\mu)A &= (\lambda+\mu)[(a_1,...,a_n)] \\ &= ((\lambda+\mu)a_1,...,(\lambda+\mu)a_n) \\ &= (\lambda a_1+\mu a_1,...,\lambda a_n+\mu a_n) \\ &= \lambda A+\mu A.\end{aligned}$$

8. Se tiene que $1 \cdot A = A \quad \forall A \in \mathbb{R}^n$.

   Demostración. $1\,(a_1,...,a_n) = (1 \cdot a_1,...,1 \cdot a_n) = (a_1,...,a_n) = A$.

Definición 37 *Un conjunto que satisface las propiedades 1 a 8 se llama espacio vectorial sobre los reales.*

Nótese que en $\mathbb{R}^n$, también se cumple que $-A = (-1)A$, ya que

$$(-1)A = (-1)(a_1,...,a_n) = ((-1)\cdot a_1,...,(-1)\cdot a_n)) = (-a_1,...,-a_n) = -A.$$

También, como $0[(a_1,...,a_n)] = (0 \cdot a_1,...,0 \cdot a_n) = \overline{0}$, en $\mathbb{R}^n$ se cumple que $0 \cdot A = \overline{0} \quad \forall A \in \mathbb{R}^n$.

Finalmente, nótese que si $\lambda A = 0$, entonces $\lambda = 0$ o $A = \overline{0}$, donde $\lambda \in \mathbb{R}$ y $A \in \mathbb{R}^n$.

Esto se tiene, ya que si $\lambda \neq 0$, como $\lambda\, A = (\lambda a_1,...,\lambda a_n) = \overline{0}$, se sigue que $\lambda\, a_i = 0 \quad \forall\, i$, por lo que $\lambda^{-1}\,\lambda\, a_i = \lambda^{-1}\, 0 = 0$, y $a_i = 0 \ \forall\, i$.

De manera análoga se pueden definir espacios vectoriales sobre otros campos, por ejemplo

$$\mathbb{C}^n = \underbrace{\mathbb{C} \times \cdots \times \mathbb{C}}_{n-veces}$$

es un espacio vectorial sobre $\mathbb{C}$ o

$$\mathbb{Q}^n = \underbrace{\mathbb{Q} \times \cdots \times \mathbb{Q}}_{n-veces}$$

lo es sobre $\mathbb{Q}$.

3.3. Subespacios vectoriales

Definición 38 *Sea $W \subset \mathbb{R}^n$, se dice que W es un subespacio vectorial de $\mathbb{R}^n$, si se cumplen las siguientes propiedades*

a) $\overline{0} \in W$,

b) $\forall\, A, B \in W$, *se tiene* $A + B \in W$,

c) $\forall\, A \in W$ *y* $\lambda \in \mathbb{R}$, *se tiene* $\lambda A \in W$.

Ejemplos

1. Sea $W = \{(x, y) \in \mathbb{R}^2 \,|\, x = 0\}$, *i.e.*, W es el eje de las ordenadas en $\mathbb{R}^2$. Entonces W es un subespacio vectorial, ya que

   a) $\overline{0} \in W$,

   b) si $A,\ B \in W$, entonces $A = (0, y_1)$ y $B = (0, y_2)$, por lo que $A + B = (0, y_1 + y_2) \in W$,

   c) si $A \in W$ *y* $\lambda \in \mathbb{R}$, entonces $A = (0, y)$ y $\lambda A = (0, \lambda y) \in W$.

2. Sea $W = \{(x, y) \in \mathbb{R}^2 \,|\, (x, y) = \lambda(x_0, y_0), \lambda \in \mathbb{R}\}$, donde $(x_0, y_0) \in \mathbb{R}^2$ es un vector fijo, entonces W es un subespacio vectorial de $\mathbb{R}^2$:

   a) $\overline{0} = 0(x_0, y_0) \qquad \therefore \overline{0} \in W$,

   b) si $A_1, A_2 \in W$, entonces $A_1 = \lambda_1(x_0, y_0)$ y $A_2 = \lambda_2(x_0, y_0)$
   $\therefore A_1 + A_2 = (\lambda_1 x_0 + \lambda_2 x_0,\ \lambda_1 y_0 + \lambda_2 y_0) = (\lambda_1 + \lambda_2)(x_0, y_0) \in W$,

   c) si $A \in W$ *y* $\mu \in \mathbb{R}$, entonces $A = \lambda(x_0, y_0)$
   $\therefore \mu A = \mu[\lambda(x_0, y_0)] = (\mu\lambda)(x_0, y_0) \in W$.

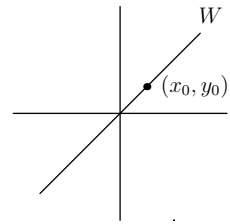


Figura 3.13: La recta $W = \{(x, y) \in \mathbb{R}^2 \,|\, (x, y) = \lambda(x_0, y_0), \lambda \in \mathbb{R}\}$ es un subespacio vectorial de $\mathbb{R}^2$

3. Sea $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1 = 0 = x_2\}$, entonces W es un subespacio vectorial de $\mathbb{R}^3$:

   a) $\overline{0} \in W$,

   b) si $(0, 0, x_3) \in W$ y $(0, 0, x_3') \in W$, se tiene $(0, 0, x_3 + x_3') \in W$,

   c) si $(0, 0, x_3) \in W$, entonces $\lambda(0, 0, x_3) = (0, 0, \lambda x_3) \in W$.

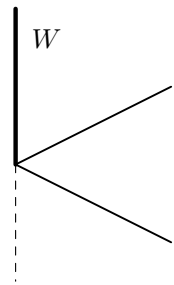


Figura 3.14: La recta $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1 = 0 = x_2\}$ es un subespacio vectorial de $\mathbb{R}^3$

4. El plano $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_3 = 0\}$ es un subespacio vectorial de $\mathbb{R}^3$, ya que $\overline{0} \in W$; $(x_1, x_2, 0)+(x_1', x_2', 0) = (x_1+x_1', x_2+x_2', 0) \in W$ y $\lambda(x_1, x_2, o) = (\lambda x_1, \lambda x_2, 0) \in W$.

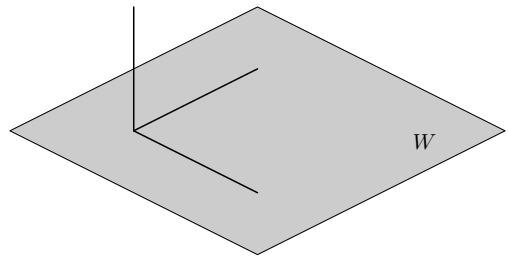


Figura 3.15: El plano $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_3 = 0\}$ es un subespacio vectorial de $\mathbb{R}^3$

5. Sean $A, B \in \mathbb{R}^3$ y $W = \{\overline{x} \in \mathbb{R}^3 \,|\, \overline{x} = \lambda A + \mu B \;\ni\; \lambda, \mu \in \mathbb{R}\}$, entonces W es un subespacio vectorial. Obsérvese que W es el plano "generado" por A y B, véase la Figura 3.16. Esto se sigue, ya que se cumplen las tres condiciones:

   a) $\overline{0} = 0 \cdot A + 0 \cdot B \quad \therefore\ \overline{0} \in W$,

   b) si $\overline{x}_1 \in W$ y $\overline{x}_2 \in W$, $\overline{x}_1 = \lambda_1 A + \mu_1 B$ y $\overline{x}_2 = \lambda_2 A + \mu_2 B$, entonces

   $$\overline{x}_1 + \overline{x}_2 = \lambda_1 A + \mu_1 B + \lambda_2 A + \mu_2 B = (\lambda_1 + \lambda_2)A + (\mu_1 + \mu_2)B \in W,$$

   c) si $\overline{x} \in W$ y $t \in \mathbb{R}$, $\overline{x} = \lambda A + \mu B$, entonces se tiene que

   $$t\,\overline{x} = t\,(\lambda A + \mu B) = (t\,\lambda)A + (t\,\mu)B \in W.$$

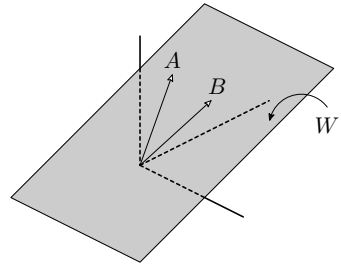


Figura 3.16: El plano por el origen W en $\mathbb{R}^3$ dado por la siguiente ecuación $\{\overline{x} \in \mathbb{R}^3 \,|\, \overline{x} = \lambda A + \mu B \;\ni\; \lambda, \mu \in \mathbb{R}\}$ es un subespacio vectorial

6. $\overline{0}$ es un subespacio vectorial $\mathbb{R}^n$:

   $\overline{0} + \overline{0} = \overline{0}$ y $\lambda\,\overline{0} = \overline{0} \quad \forall\, \lambda \in \mathbb{R}$.

Se define el producto punto de dos vectores en $\mathbb{R}^3$, como sigue

$$(x_1, y_1, z_1) \cdot (x_2, y_2, z_2) = x_1\, x_2 + y_1\, y_2 + z_1\, z_2.$$

Análogamente se define en $\mathbb{R}^2$ (o en $\mathbb{R}^n$). Este producto determina la perpendicularidad u ortogonalidad, como se muestra a continuación.

Definición 39 *Los vectores A y B en $\mathbb{R}^2$ (o $\mathbb{R}^3$) son ortogonales o perpendiculares, si $|A - B| = |A + B|$, i.e., las magnitudes de sus diagonales son iguales.*

Teorema 3.3.1 *Dos vectores A y B en $\mathbb{R}^2$ (o $\mathbb{R}^3$) son perpendiculares si y sólo si $A \cdot B = 0$.*

DEMOSTRACIÓN. $|A+B|^2 - |A-B|^2 = 4\, A \cdot B.$

□

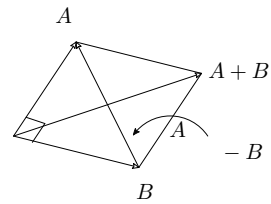


Figura 3.17: Una manera de definir que dos vectores sean perpendiculares es que sus diagonales midan lo mismo

Ahora, si se tienen dos vectores $A, B \in \mathbb{R}^3$ de longitud unitaria, como en la Figura 3.18, aplicando el Teorema 3.3.1, se sigue que el vector $A - tB$ es perpendicular al vector B si y sólo si $(A - tB) \cdot B = 0$. Esta condición es equivalente a su vez a que $A \cdot B = t$. Por otra parte, dicha ortogonalidad se tiene si $t = \cos\theta$, véase la Figura 3.18. En consecuencia

$$A \cdot B = \cos\theta,$$

por supuesto, $A \cdot B = 0 \iff \theta = \pi/2$.

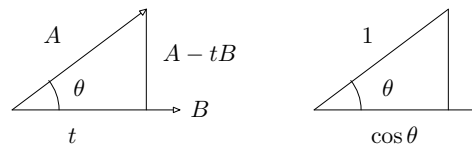


Figura 3.18: Si $|A| = 1$, $|B| = 1$, $t = \cos\theta$, entonces $A - t\,B$ es perpendicular a B.

Ahora, la ecuación de un plano en $\mathbb{R}^3$ está dada por

$$(P - P_0) \cdot n = 0,$$

es decir,

$$[(x, y, z) - (x_0, y_0, z_0)] \cdot (n_1, n_2, n_3) = 0,$$

donde $n = (n_1, n_2, n_3)$ es una normal (es decir, perpendicular) al plano y $P_0 = (x_0, y_0, z_0)$ es un punto del plano. Dicho de otra manera, la ecuación general de un plano en $\mathbb{R}^3$ está dada por

$$ax + by + cz = d.$$

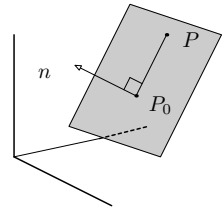


Figura 3.19: Ecuación general de un plano P en $\mathbb{R}^3$
$p \cdot n =$ constante, donde n es la normal y $p \in P$

Otros ejemplos

1. Probamos que $W = \{(x, y) \in \mathbb{R}^2 \,|\, 2x - 3y = 0\}$ es un subespacio vectorial de $\mathbb{R}^2$:

   a) $2 \cdot 0 - 3 \cdot 0 = 0 \qquad \therefore \quad 0 \in W$,

   b) si (x_1, y_1) y $(x_2, y_2) \in W$, se tiene $2x_1 - 3y_1 = 0$ y $2x_2 - 3y_2 = 0$
   $\therefore\ 2(x_1 + x_2) - 3(y_1 + y_2) = 0 \ \therefore\ (x_1 + x_2, y_1 + y_2) \in W$,

   c) si $(x, y) \in W$ y $\lambda \in \mathbb{R}$, $2x - 3y = 0 \quad \therefore \quad 2\lambda x - 3\lambda y = 0$
   $\therefore\ (\lambda x, \lambda y) \in W$.

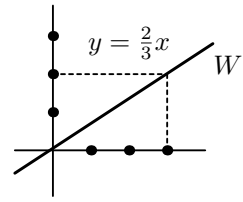


Figura 3.20: La recta $W = \{(x, y) \in \mathbb{R}^2 \,|\, 2x - 3y = 0\}$ es un subespacio vectorial

2. Probamos que

$$W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + y + z = 0,\ 3x - 5y + z = 0\}$$

es un subespacio vectorial en $\mathbb{R}^3$. Nótese que W es la intersección de 2 planos distintos, por lo que es una recta. Como la intersección de subespacios vectoriales es un subespacio vectorial (ejercicio), basta probar que

$$W_1 = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + y + z = 0\}$$

y

$$W_2 = \{(x, y, z) \in \mathbb{R}^3 \mid 3x - 5y + z = 0\}$$

lo son. Probamos que W_1 es un subespacio vectorial y la prueba para W_2 es análoga.

a) Como $2 \cdot 0 + 0 + 0 = 0, \qquad \overline{0} \in W_1,$

b) si

$$\begin{array}{rcrcrcl} 2x_1 & + & y_1 & + & z_1 & = & 0, \\ 2x_2 & + & y_2 & + & z_2 & = & 0 \end{array}$$

entonces

$$2(x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) = 0,$$

i.e., la suma de vectores en W_1 está en W_1,

c) si $2x + y + z = 0$ y $\lambda \in \mathbb{R}$, entonces $2\lambda x + \lambda y + \lambda z = 0$ y $(\lambda x, \lambda y, \lambda z) \in W_1$.

EJERCICIOS 3.3

1. Pruebe que un plano en $\mathbb{R}^3$ por el origen, *i.e.*, un conjunto de la forma $W = \{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz = 0,\ a, b, c \in \mathbb{R}\}$ es un subespacio vectorial.

2. Pruebe que $W = \{(x, y, z) \in \mathbb{R}^3 \mid 3x - 3y + z = 0\}$ es un subespacio vectorial de $\mathbb{R}^3$.

3. Pruebe que $W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_2 + x_3 + x_4 = 0\}$ es un subespacio vectorial.

4. Pruebe que la intersección de 2 subespacios vectoriales en $\mathbb{R}^n$ es un subespacio vectorial.

3.4. Combinaciones lineales, dependencia e independencia lineal

Definición 40 *Sean* $A_1,\ A_2\ ,...,\ A_k$ *vectores en* $\mathbb{R}^n$, *se dice que* $C \in \mathbb{R}^n$ *es combinación lineal de* $A_1,\ A_2\ ,...,\ A_k$, *si existen* $\lambda_1,\ \lambda_2\ ,...,\ \lambda_k \in \mathbb{R}$ *tales que* $C = \lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k$.

Ejemplos

1. El vector $(-4, 5)$ en $\mathbb{R}^2$ es combinación lineal de $A_1 = (1, 1)$ y $A_2 = (2, -1)$, ya que $(-4, 5) = 2(1, 1) + (-3)(2, -1)$.

2. En $\mathbb{R}^2$ $(2, 1)$ no es combinación lineal de los vectores $(1, 0)$ y $(-2, 0)$, esto se sigue ya que $\lambda(1, 0) + \mu(-2, 0) = (\lambda - 2\,\mu, 0)$.

3. En $\mathbb{R}^3$ todo vector es combinación lineal de los vectores $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$:

$$(x, y, z) \ = \ x\ (1, 0, 0) \ + \ y(0, 1, 0) \ + \ z(0, 0, 1).$$

4. En $\mathbb{R}^n$ todo vector es combinación lineal de los vectores canónicos $e_1, e_2, ..., e_n$, donde $e_j = (0, 0, ..., \underbrace{0, 1, 0}_{\text{j-ésimo lugar}}, ..., 0)$:

$$(x_1, x_2, ..., x_n) \ = \ x_1 e_1 \ + \ x_2 e_2 \ + \ \cdots \ + \ x_n e_n.$$

5. Dados vectores $A_1, A_2, ..., A_k$, cualquiera de ellos es combinación lineal del conjunto de todos ellos, por ejemplo

$$A_1 \ = \ 1 \cdot A_1 \ + \ 0 \cdot A_2 \ + \ ... \ + \ 0 \cdot A_k.$$

6. $\overline{0}$ es combinación lineal de cualquier conjunto de vectores $A_1, A_2, ..., A_k$:

$$\overline{0} \ = \ 0 \cdot A_1 \ + \ 0 \cdot A_2 \ + \ \cdots \ + \ 0 \cdot A_k.$$

7. Si tanto C como D son combinación lineal de $A_1, ..., A_k$, entonces $C + D$ también lo es:

$$C \ = \ \sum_{i=1}^{k} \lambda_i A_i \ , \ D \ = \ \sum_{i=1}^{k} \mu_i A_i \quad \therefore \ C \ + \ D \ = \ \sum_{i=1}^{k} (\lambda_i + \mu_i)\, A_i.$$

8. En $\mathbb{R}^3$ todo vector (x, y, z) es combinación lineal de $(1, 0, 0), (1, 1, 0)$ y $(1, 1, 1)$.

   Para probar este hecho se toma $(x, y, z) \in \mathbb{R}^3$, se buscan escalares $\lambda_1, \lambda_2, \lambda_3$ tales que

$$\begin{aligned}(x, y, z) &= \lambda_1(1, 0, 0) + \lambda_2(1, 1, 0) + \lambda_3(1, 1, 1)\\ &= (\lambda_1 + \lambda_2 + \lambda_3, \lambda_2 + \lambda_3, \lambda_3),\end{aligned}$$

   tomando $\lambda_3 = z$, se tiene $\lambda_2 = y - \lambda_3 = y - z$, y $x = \lambda_1 + y - z + z$, *i.e.*, $\lambda_1 = x - y$.

9. Se expresa al vector $(8, -1)$ como combinación lineal de $(2, 1)$ y $(3, -1)$.

   Para esto, se buscan λ y μ tales que $(8, -1) = \lambda(2, 1) + \mu(3, -1)$, *i.e.*,

$$\begin{cases} 8 = 2\lambda + 3\mu \\ -1 = \lambda - \mu \end{cases} \Leftrightarrow \begin{cases} 8 = 2\lambda + 3\mu \\ -3 = 3\lambda - 3\mu, \end{cases}$$

   sumando tenemos $5 = 5\lambda$, $\lambda = 1$ y $\mu = 2$.

Definición 41 *Dados vectores $A_1, ..., A_k$ en $\mathbb{R}^n$, al conjunto de combinaciones lineales de $A_1, A_2, ..., A_k$, es decir,*

$$\{\lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k \,|\, \lambda_i \in \mathbb{R}, \forall i\}.$$

se le llama el subespacio generado por $A_1, A_2, ..., A_k$, y se le denota por $< A_1, A_2, ..., A_k >$.

Proposición 3.4.1 *Dados vectores $A_1, ..., A_k$, entonces*

$$W =< A_1, A_2, ..., A_k >$$

es un subespacio vectorial.

Demostración. El vector nulo está en W, ya que $\overline{0} = 0 \cdot A_1 + \cdots + 0 \cdot A_k$. Ahora, si $\lambda_1 A_1 + \cdots + \lambda_k A_k$, $\mu_1 A_1 + \cdots + \mu_k A_k \in W$, entonces

$$(\lambda_1 + \mu_1)A_1 + \cdots + (\lambda_k + \mu_k)A_k \in W,$$

y si $\lambda \in \mathbb{R}$, se tiene

$$\lambda \lambda_1 A_1 + \cdots + \lambda \lambda_k A_k \in W.$$

□

Dependencia lineal

Definición 42 *Se dice que* $C \in \mathbb{R}^n$ *depende linealmente de* $A_1, A_2, ..., A_k$, *si* C *es combinación lineal de* $A_1, A_2, ..., A_k$, *i.e., si existen* $\lambda_1, \lambda_2, ..., \lambda_k \in \mathbb{R}$ *tales que*

$$C = \lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k.$$

Obsérvese que C depende linealmente de $A_1, A_2, \ldots, A_k$, si y sólo si $C \in < A_1, ..., A_k >$.

Definición 43 *Se dice que los vectores* $A_1, A_2, ..., A_k$ *en* $\mathbb{R}^n$ *son linealmente dependientes, si alguno de ellos depende linealmente de los otros.*

Por ejemplo, si $A = (1, 1, 0)$, $B = (4, 4, 0)$ y $C = (0, 0, 7)$, entonces A, B, C son linealmente dependientes, ya que $B = 4 \cdot A + 0 \cdot C$.

Obsérvese que cuando se dice que uno de ellos depende linealmente de los otros, no se tiene que necesariamente cada uno de ellos es combinación lineal de los otros, por ejemplo si A, B y C son como arriba, C no es combinación lineal de A y B.

Teorema 3.4.2 *Sean* $A_1, A_2, ..., A_k$ *vectores en* $\mathbb{R}^n$, *entonces estos vectores son linealmente dependientes si y sólo si existe una combinación lineal de ellos igual a* $\overline{0}$ *con algún coeficiente distinto de 0, i.e.,*

$$\lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k = \overline{0}$$

y $\lambda_j \neq 0$ *para alguna* $j \in \{1, 2, ..., k\}$.

Demostración.
$\Rightarrow$) Si $A_1, A_2, ..., A_k$ son linealmente dependientes, alguno de ellos, digamos A_j, depende linealmente de los otros, *i.e.*, existen $\lambda_1, \lambda_2, ..., \lambda_{j-1}, \lambda_{j+1}, ..., \lambda_k$ números reales, tales que

$$A_j = \sum_{i=1}^{j-1} \lambda_i A_i + \sum_{i=j+1}^{k} \lambda_i A_i,$$

y por lo tanto

$$\lambda_1 A_1 + \cdots + \lambda_{j-1} A_{j-1} + (-1) A_j + \lambda_{j+1} A_{j+1} + \cdots + \lambda_k A_k = \overline{0}.$$

$\Leftarrow$) Viceversa, si existen $\lambda_1, \lambda_2, ..., \lambda_k \in \mathbb{R} \ni \lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k = \overline{0}$, donde $\lambda_j \neq 0$, entonces despejando se tiene

$$A_j = \left(\frac{-\lambda_1}{\lambda_j} A_1\right) + \cdots + \left(\frac{-\lambda_{j-1}}{\lambda_j} A_{j-1}\right) + \left(\frac{-\lambda_{j+1}}{\lambda_j} A_{j+1}\right) + \cdots + \left(\frac{-\lambda_k}{\lambda_j} A_k\right).$$

□

Definición 44 *Se dice que una colección de vectores $A_1, A_2, ..., A_k$ en $\mathbb{R}^n$ son linealmente independientes, si no son linealmente dependientes.*

Esta definición se puede reformular como sigue: $A_1, A_2, ..., A_k$ son linealmente independientes, si ninguno de ellos es combinación lineal de los otros.

Corolario 3.4.3 *Sean $A_1, A_2, ..., A_k$ vectores en $\mathbb{R}^n$. Entonces, estos vectores son linealmente independientes si y sólo si para toda combinación lineal $\lambda_1 A_1 + \cdots + \lambda_k A_k = \overline{0}$, se tiene que $\lambda_i = 0 \ \forall i$.*

Demostración.

$\Rightarrow$) Si $A_1, A_2, ..., A_k$ son linealmente independientes, entonces no son linealmente dependientes. En consecuencia, se sigue del Teorema 3.4.2 que si

$$\lambda_1 A_1 + \cdots + \lambda_k A_k = \overline{0},$$

necesariamente $\lambda_i = 0 \ \forall i$.

$\Leftarrow$) Si para toda combinación lineal $\lambda_1 A_1 + \cdots + \lambda_k A_k = \overline{0}$, se tiene que $\lambda_i = 0 \ \ \forall i$, se sigue del Teorema 3.4.2 que $A_1, A_2, ..., A_k$ no pueden ser linealmente dependientes *i.e.*, son linealmente independientes.

□

Por ejemplo, en $\mathbb{R}^2$ los vectores $(1,0)$ y $(1,1)$ son linealmente independientes, en efecto si $\lambda(1,0) + \mu(1,1) = \overline{0}$, entonces $\lambda + \mu = 0$ y $\mu = 0$, por lo que $\lambda = 0$.

Nótese que si un conjunto de vectores $\{A_1, A_2, ..., A_k\}$ es linealmente dependiente, entonces cualquier otro conjunto que lo contenga, por ejemplo $\{A_1, A_2, \ldots, A_k, A_{k+1}, \ldots, A_s\}$ también es linealmente dependiente.

Esto se sigue ya que existen $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$, alguno distinto de 0, tales que $\lambda_1 A_1 + \cdots + \lambda_k A_k = 0$, por lo cual se tiene

$$\lambda_1 A_1 + \cdots + \lambda_k A_k + 0 \cdot A_{k+1} + \ldots + 0 \cdot A_{k+s} = 0.$$

Ejemplos

1. Probamos que los vectores $D_1 = (1,0,0)$, $D_2 = (1,1,0)$ y $D_3 = (1,1,1)$ son linealmente independientes en $\mathbb{R}^3$:

   si $\lambda_1 D_1 + \lambda_2 D_2 + \lambda_3 D_3 = \overline{0}$, entonces

$$\begin{array}{rcrcrcl} \lambda_1 & + & \lambda_2 & + & \lambda_3 & = & 0, \\ & & \lambda_2 & + & \lambda_3 & = & 0, \\ & & & & \lambda_3 & = & 0, \end{array}$$

   y por lo tanto $\lambda_1, \lambda_2, \lambda_3 = 0$.

2. En $\mathbb{R}^n$, los vectores $D_1 = (1, 0, \ldots, 0)$, $D_2 = (1, 1, 0, \ldots, 0)$, $\ldots$, $D_i = (1, \ldots, \underbrace{1, 1}_{\text{i-ésima coordenada}}, 0, \ldots, 0)$, $\ldots$, $D_n = (1, \ldots, 1)$

   son linealmente independientes:

$$\text{si} \qquad \sum_{i=1}^{n} \lambda_i \, D_i = \overline{0},$$

   se tiene

$$\begin{array}{rcrcrcrcl} \lambda_1 & + & \lambda_2 & + & \cdots & + & \lambda_n & = & 0, \\ & & & & \vdots & & & & \\ & & \lambda_2 & + & \cdots & + & \lambda_n & = & 0, \\ & & & & \vdots & & & & \\ & & & & & & \lambda_n & = & 0, \end{array}$$

   y por lo tanto $\lambda_1, \lambda_2, \ldots, \lambda_n = 0$.

3. Probamos que en $\mathbb{R}^2$ los vectores $(-1,1), (2,3)$ y $(-5,2)$ son linealmente dependientes:

   si $\lambda_1(-1,1) + \lambda_2(2,3) + \lambda_3(-5,2) = \overline{0}$, entonces

$$\begin{cases} -\lambda_1 + 2\lambda_2 - 5\lambda_3 = 0, \\ \lambda_1 + 3\lambda_2 + 2\lambda_3 = 0, \end{cases}$$

   sumando se tiene $5\lambda_2 - 3\lambda_3 = 0$, tomando $\lambda_2 = 3$ se obtiene $\lambda_3 = 5$. Sustituyendo estos valores, por ejemplo, en la primera ecuación se tiene $-\lambda_1 + 6 - 25 = 0$ y $\lambda_1 = -19$. Se concluye que los vectores son linealmente dependientes, ya que $-19(-1,1) + 3(2,3) + 5(-5,2) = \overline{0}$.

EJERCICIOS 3.4
1. Demuestre que en $\mathbb{R}^2$ todo vector es combinación lineal de $(1,2)$ y $(0,1)$.
2. Exprese el vector $(3,4)$ como combinación lineal de $(1,2)$ y $(0,1)$.
3. Dados $u_1 = (3,-1,0)$, $u_2 = (0,1,0)$ y $u_3 = (5,4,0)$, exprese el vector u_3 como combinación lineal de los otros dos.
4. Probar que en $\mathbb{R}^3$ los vectores $(2,0,0), (1,3,0)$ y $(1,2,4)$ son linealmente independientes.
5. Probar que en $\mathbb{R}^3$ los vectores $(1,0,-2), (1,1,0), (1,0,4)$ y $(1,1,1)$ son linealmente dependientes.

3.5. Bases

Definición 45 *Un conjunto* $\{A_1, A_2, ..., A_r\}$ *de vectores en* $\mathbb{R}^n$ *se llama base del subespacio vectorial* $W \subset \mathbb{R}^n$, *si*

a) $\{A_1, A_2, ..., A_r\}$ *son linealmente independientes,*

b) $< A_1, A_2, ..., A_r >= W$, *i.e.* $A_1, A_2, ..., A_r$ *generan* W.

Ejemplos

1. Si $e_1 = (1,0)$ y $e_2 = (0,1)$, entonces $\{e_1, e_2\}$ es una base de $\mathbb{R}^2$, esto se sigue ya que e_1 y e_2 son linealmente independientes y generan $\mathbb{R}^2$: $(x,y) = x\,e_1 + y\,e_2$.

2. Si $D_1 = (1,0)$ y $D_2 = (1,1)$, entonces $\{D_1, D_2\}$ es una base de $\mathbb{R}^2$: si $\lambda_1(1,0) + \lambda_2(1,1) = 0$, entonces $\lambda_1 = \lambda_2 = 0$ y D_1, D_2 son linealmente independientes. Ahora probamos que generan: dado $(x,y) \in \mathbb{R}^2$, se debe resolver el sistema $(x,y) = \lambda_1(1,0) + \lambda_2(1,1)$, es decir,

$$\begin{cases} x = \lambda_1 + \lambda_2, \\ y = \lambda_2. \end{cases}$$

   El cual evidentemente tiene solución.

3. En $\mathbb{R}^n$ los vectores canónicos $e_1, e_2, ..., e_n$, forman una base, donde

$$e_i = (0,0,...,0,1,0,...,0),$$

con coordenada i-ésima igual a 1 y las otras 0. Estos son linealmente independientes ya que si $\sum_{i=1}^{n} \lambda_i e_i = 0$, se tiene $\lambda_i = 0 \ \ \forall\ i$. Por otra parte, se mostró en la sección anterior que todo vector en $\mathbb{R}^n$ es combinación lineal de éstos: si $\overline{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, $\overline{x} = \sum_{i=1}^{n} x_i e_i$.

4. También los vectores $D_1, D_2, ..., D_n$ forman una base de $\mathbb{R}^n$, donde $D_1 = (1, 0, 0, ..., 0), D_2 = (1, 1, 0, ..., 0)$,...,$D_n = (1, 1, 1, ..., 1)$.
   Se mostró al final de la sección anterior el hecho de que estos vectores son linealmente independientes, mostramos ahora que generan $\mathbb{R}^n$. Dado $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$ se deben encontrar $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$ ∋

$$\sum_{i=1}^{n} \lambda_i D_i = (x_1, ..., x_n),$$

   esto es,

$$\begin{cases} \lambda_1 + \lambda_2 + \cdots + \lambda_n = x_1, \\ \lambda_2 + \cdots + \lambda_n = x_2, \\ \vdots = \vdots \\ \lambda_n = x_n, \end{cases}$$

   y la solución está dada por: $\lambda_n = x_n$,
   $\lambda_{n-1} + \lambda_n = x_{n-1}$, *i.e.*, $\lambda_{n-1} = x_{n-1} - x_n$,
   $\lambda_{n-2} = x_{n-2} - (x_{n-1} - x_n) - x_n = x_{n-2} - x_{n-1}$,
   $\lambda_{n-3} = x_{n-3} - (x_{n-2} - x_{n-1}) - (x_{n-1} - x_n) - x_n = x_{n-3} - x_{n-2}$,
   inductivamente se sigue que $\lambda_{n-j} = x_{n-j} - x_{n-j+1}$ (ejercicio).

5. Sea $W = \{(x, y, z) \in \mathbb{R}^3 \,|\, 2x + 3y - z = 0\}$, W es un subespacio vectorial con base $\{(1, 0, 2), (0, 1, 3)\}$. Por una parte los vectores son linealmente independientes: $\lambda(1, 0, 2) + \mu(0, 1, 3) = \overline{0} \Rightarrow \lambda = \mu = 0$; y por otra estos vectores también generan W: dado $(x, y, z) \in W$, se tiene $z = 2x + 3y$, hay que probar
   $$(x, y, z) = \lambda(1, 0, 2) + \mu(0, 1, 3),$$
   i.e., $(x, y, z) = (\lambda, \mu, 2\lambda + 3\mu)$, tomando $\lambda = x$ y $\mu = y$ se sigue el resultado.

6. Sea $A \in \mathbb{R}^n - \{\overline{0}\}$ y $W = \{\overline{x} \in \mathbb{R}^n \,|\, \overline{x} = \lambda A, \lambda \in \mathbb{R}\}$, entonces A es una base de W: A es linealmente independiente ya que si $\lambda A = \overline{0}$, entonces $\lambda = 0$, también genera ya que si $\overline{y} \in W$, entonces $\overline{y} = \lambda A$.

7. En $\mathbb{R}^3$, $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_3 = 0\}$ es un subespacio vectorial con base $e_1 = (1, 0, 0)$ y $e_2 = (0, 1, 0)$: ya se mostró que e_1 y e_2 son linealmente independientes, ahora dado $(x, y, z) \in W$ se tiene

$$(x, y, z) = xe_1 + ye_2.$$

Geométricamente es evidente que si A, B, C son vectores en $\mathbb{R}^2$ (veáse la Figura 3.21), éstos deben ser linealmente dependientes; probaremos esto de manera formal y más general.

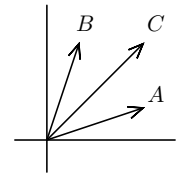


Figura 3.21: En $\mathbb{R}^2$ cualesquiera tres vectores son linealmente dependientes

Teorema 3.5.1 *Sea W un subespacio vectorial generado por k vectores, entonces cualquier conjunto de $k+1$ vectores en W es linealmente dependiente. Y por ende, cualquier otro conjunto de $k+m$ vectores en W, $m \geq 1$, también es linealmente dependiente.*

Antes de probar el teorema mostramos un ejemplo. Sea W un subespacio vectorial generado por A_1 y A_2 en $\mathbb{R}^n$ y

$$B_1 = \alpha_{11}A_1 + \gamma_1 A_2,$$

$$B_2 = \alpha_{21}A_1 + \gamma_2 A_2,$$

$$B_3 = \alpha_{31}A_1 + \gamma_3 A_2,$$

donde alguna $\gamma_i \neq 0$ (si $\gamma_i = 0 \ \forall\, i$ son tres vectores generados por un solo vector, lo cual no es el planteamiento), digamos $\gamma_1 \neq 0$, entonces

$$B_2' = B_2 - \tfrac{\gamma_2}{\gamma_1} B_1 = (escalar) A_1,$$

$$B_3' = B_3 - \tfrac{\gamma_3}{\gamma_1} B_1 = (escalar) A_1,$$

por lo que B_2' y B_3' son linealmente dependientes. En consecuencia, $B_1, B_2,$ y B_3 también lo son, ya que una combinación lineal de B_2' y B_3' igual a $\overline{0}$, con algún coeficiente distinto de 0, genera una análoga para B_1, B_2, B_3 (el coeficiente de B_2 es el mismo que el de B_2', y también el de B_3 es el de B_3').

Demostración del Teorema 3.5.1. Se prueba por inducción sobre k.

Si $k = 1$, $W = \{\lambda A \,|\, \lambda \in \mathbb{R}\}$, donde $A \in \mathbb{R}^n - \{0\}$ es un vector fijo. Dados B_1 y $B_2 \in W$, $B_1 \neq B_2$, entonces $B_1 = \lambda_1 A$ y $B_2 = \lambda_2 A$, donde podemos suponer $\lambda_1 \neq 0 \therefore B_2 = \lambda_2 \left(\frac{1}{\lambda_1} B_1\right)$ y B_1, B_2 son linealmente dependientes.

Suponiendo cierto el resultado para k, ahora probamos para $k+1$. Sean $A_1, A_2, ..., A_{k+1}$ vectores que generan W y $B_1, B_2, ..., B_{k+2}$ vectores en W, entonces

$$\begin{array}{ccccccccccc}
B_1 & = & \alpha_{11}A_1 & + & \cdots & + & \alpha_{1k}A_k & + & \gamma_1 A_{k+1} \\
B_2 & = & \alpha_{21}A_1 & + & \cdots & + & \alpha_{2k}A_k & + & \gamma_2 A_{k+1} \\
\vdots & & & & & & & \vdots & \\
B_{k+2} & = & \alpha_{(k+2)1}A_1 & + & \cdots & + & \alpha_{(k+2)k}A_k & + & \gamma_{(k+2)} A_{k+1}.
\end{array}$$

Se tienen 2 casos:

Caso 1: alguna $\gamma_i \neq 0$. Sin perder generalidad, se puede tomar $\gamma_1 \neq 0$. En este caso los vectores $B_2', ..., B_{k+2}'$, definidos por

$$B_2' = B_2 - \frac{\gamma_2}{\gamma_1} B_1, \ \ldots, B_{k+2}' = B_{k+2} - \frac{\gamma_{k+2}}{\gamma_1} B_1,$$

son $k+1$ vectores que pertenecen al subespacio vectorial generado por $A_1, A_2, ..., A_k$ (análogamente al ejemplo, se *eliminó* A_{k+1} y B_1).

Se sigue entonces de la hipótesis de inducción que $B_2', ..., B_{k+2}'$ son linealmente dependientes, por lo que existe una combinación lineal

$$\lambda_2 B_2' + \cdots + \lambda_{k+2} B_{k+2}' = \overline{0}$$

con alguna $\lambda_i \neq 0$, sustituyendo se tiene

$$\lambda_2 B_2 + \cdots + \lambda_{k+2} B_{k+2} + (coeficiente) B_1 = \overline{0}$$

$\therefore$ $B_1, B_2, ..., B_{k+2}$ son linealmente dependientes.

Caso 2: $\gamma_i = 0 \ \forall \, i$. En este caso, $B_1, B_2, ..., B_{k+2}$ están en el subespacio generado por $A_1, ..., A_k$, se sigue entonces de la hipótesis de inducción que $B_1, B_2, ..., B_{k+1}$ son linealmente dependientes, y por lo tanto $B_1, B_2, ..., B_{k+2}$ también lo son. □

Corolario 3.5.2 *En* $\mathbb{R}^n$, *cualquier conjunto de más de n vectores es linealmente dependiente.*

DEMOSTRACIÓN. La base canónica $e_1, e_2, ..., e_n$ genera $\mathbb{R}^n$. □

EXISTENCIA DE BASES

Lema 3.5.3 *Sea* W *un subespacio vectorial de* $\mathbb{R}^n$ *generado por los vectores linealmente independientes* $B_1, B_2, ..., B_k$. *Supóngase también que* $B \in \mathbb{R}^n - W$, *entonces* $B, B_1, B_2, ..., B_k$ *son linealmente independientes.*

DEMOSTRACIÓN. Sea

$$\lambda_1 B_1 + \cdots + \lambda_k B_k + \lambda B = \overline{0}$$

una combinación lineal. Necesariamente $\lambda = 0$, ya que de otra manera $B \in W$, por lo cual

$$\lambda_1 B_1 + ... + \lambda_k B_k = \overline{0}.$$

Se sigue entonces que $\lambda_i = 0 \ \ \forall \ i$ (ya que $B_1, B_2, ..., B_k$ son linealmente independientes),

$$\therefore \ \{B, B_1, B_2, ..., B_k\} \quad \text{son linealmente independientes.}$$

□

Teorema 3.5.4 *Sea* W *un subespacio vectorial de* $\mathbb{R}^n$ *y* $B_1, B_2, ..., B_k$ *vectores de* W *linealmente independientes, entonces existen vectores*

$$B_{k+1}, B_{k+2}, ..., B_{k+t}$$

tales que

$$B_1, B_2, ..., B_k, ..., B_{k+t}$$

es una base de W.

Demostración. Si $B_1, B_2, ..., B_k$ generan W no hay nada que probar. De otra manera, sea $W_1 =< B_1, B_2, ..., B_k >$ y $B_{k+1} \in W - W_1$, por el Lema 3.5.3

$$\{B_1, B_2, ..., B_k, B_{k+1}\}$$

son linealmente independientes. Si estos generan W, ya terminamos, de otra manera, sea $B_{k+2} \in W - W_2$, donde $W_2 =< B_1, ..., B_{k+1} >$, etcétera.

El proceso debe terminar antes de que $k + t$ sea menor o igual a n, en caso contrario se tendrían más de n vectores linealmente independientes en $\mathbb{R}^n$, lo cual contradice el Corolario 3.5.2.

$\therefore$ Existe $t \in \mathbb{N}$ ∋ $k + t \leq n$ y los vectores $\{B_1, B_2, ..., B_{k+t}\}$ forman una base de W.

□

Corolario 3.5.5 *Todo subespacio vectorial de $\mathbb{R}^n$ (distinto de $\overline{0}$) tiene una base.*

Demostración. Si $W \neq \overline{0}$, entonces existe $A \in W$ ∋ $A \neq 0$, como $\{A\}$ es linealmente independiente, el resultado se sigue del Teorema 3.5.4.

□

dimensión

Teorema 3.5.6 *Todas las bases de un subespacio vectorial W de $\mathbb{R}^n$ tienen el mismo número de elementos.*

Demostración. Sean $\{A_1, ..., A_k\}$ y $\{B_1, ..., B_t\}$ dos bases de W, hay que probar que $k = t$.

Como $A_1, A_2, ..., A_k$ generan W y $B_1, B_2, ..., B_t$ son linealmente independientes, se sigue del Teorema 3.5.1 que $t \leq k$ (si $t > k$ dicho teorema dice que $B_1, B_2, ..., B_t$ son linealmente dependientes).

Aplicando el mismo argumento al revés se tiene $k \leq t$.

□

Definición 46 *Sea W un subespacio vectorial de $\mathbb{R}^n$, se define su dimensión como el número de elementos de cualquier base.*

Ejemplos

1. $\mathbb{R}$ tiene dimensión 1, $\mathbb{R}^2$ dimensión dos, $\mathbb{R}^3$ dimensión tres y $\mathbb{R}^n$ dimensión n.

2. Cualquier recta por el origen en $\mathbb{R}^2$ tiene dimensión 1. Una recta con estas características es de la forma

$$L = \{\lambda(x_0, y_0) \,|\, \lambda \in \mathbb{R}\},$$

donde $(x_0, y_0) \in L - \{\overline{0}\}$ es un vector fijo.

3. Si se toma el vacío como base de $\{\overline{0}\}$, entonces $\{\overline{0}\}$ tiene dimensión 0 (el número de elementos de la base es 0). De otra manera, se puede convenir que su dimensión es 0.

4. En $\mathbb{R}^3$ las rectas por el origen tienen dimensión 1 y los planos dimensión 2: si P es un plano por el origen en $\mathbb{R}^3$, sea $A_1 \neq 0$, $A_1 \in P$, debe existir también $A_2 \in P - < A_1 >$, ya que de otra manera P es una recta. Se afirma que $P = < A_1, A_2 >$; esto se deriva ya que si $A_3 \in P - < A_1, A_2 >$, se sigue del Lema 3.5.3 que A_1, A_2 y A_3 son linealmente independientes, y entonces $P = \mathbb{R}^3$ (de otra manera existe $A_4 \in \mathbb{R}^3 - < A_1, A_2, A_3 >$, y $\mathbb{R}^3$, tendría dimensión mayor a 3).

5. En $\mathbb{R}^n$ se les llama rectas por el origen a los subespacios de dimensión 1, planos por el origen a los subespacios de dimensión 2, hiperplanos a los subespacios de dimensión k, $2 < k < n$, y subespacios de codimensión 1, a los de dimensión $n - 1$ en $\mathbb{R}^n$.

EJERCICIOS 3.5

1. Acabar los argumentos formales en el ejemplo 4 antes del Teorema 3.5.1.

2. Sea $W = \{(x, y, z) \in \mathbb{R}^3 \,|\, 5x + y + z = 0\}$, encuentre una base de W y pruebe que en efecto lo es.

3. Sea $W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,|\, x_1 + x_2 + x_3 + x_4 = 0\}$, encuentre una base de W, verifique formalmente.

4. Encuentre una base para la recta,

$$W = \{(x, y, z) \in \mathbb{R}^3 \,|\, 2x - y - z = 0,\ 3x + y - 2z = 0\}.$$

Justifique rigurosamente sus afirmaciones.

5. Encuentre una base para el plano

$$W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,|\, 2\,x_1 + x_2 - x_3 + x_4 = 0,\ x_1 - x_2 - x_3 + 5x_4 = 0\}.$$

Verifique formalmente.

6. Pruebe que $(1, 2)$, $(2, -1)$ forman una base de $\mathbb{R}^2$.

Capítulo 4

Matrices y determinantes

4.1. Definiciones

En las matemáticas aparecen frecuentemente arreglos rectangulares de números. Por ejemplo, al estudiar un sistema de ecuaciones lineales

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

es natural considerar el siguiente arreglo de números

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{4.1}$$

Este tipo de arreglos se llaman matrices, las hileras verticales se llaman columnas y las horizontales renglones, en este caso se tienen m renglones y n columnas, y se dice que se trata de una matriz de $m \times n$. También con los términos constantes del sistema se obtiene una matriz de $m \times 1$.

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Por ejemplo, al sistema

$$\begin{cases} 3x & - & 2y & + & z & = & 0 \\ & & y & - & 3z & = & 7 \\ 2x & + & 7y & & & = & 4, \end{cases}$$

se le asocian las matrices

$$\begin{pmatrix} 3 & -2 & 1 \\ 0 & 1 & -3 \\ 2 & 7 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 0 \\ 7 \\ 4 \end{pmatrix}$$

Los elementos que forman una matriz se toman en un anillo como $\mathbb{Z}$, o en un campo como en $\mathbb{Q}, \mathbb{R}$ o $\mathbb{C}$. El símbolo a_{ij} denotará el número ubicado en el renglón i y la columna j. Por ejemplo en la matriz de 3×3 descrita anteriormente $a_{2\,3} = -3$. En una matriz de $n \times n$, denominada *cuadrada*, a los elementos de la forma a_{jj}, se les llama *diagonales*. Una matriz cuadrada se llama *diagonal* si sus únicos elementos distintos de 0 son los diagonales, por ejemplo

$$\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} 5 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Sean a_{ij}, las entradas de una matriz cuadrada $1 \leq i \leq n$, y $1 \leq j \leq n$. Nótese que si $i > j, a_{ij}$ está abajo de la diagonal, y si $i < j$, entonces a_{ij} está arriba de la diagonal. Para constatar este hecho considérese el renglón $i \geq 2$ y la primera columna, esto es abajo de la diagonal, moviéndose de columna en el mismo renglón, el número de columna crece, en la diagonal es igual a i, etcétera.

Definición 47 *Una matriz cuadrada se le llama triangular si todas las entradas de arriba (o todas las de abajo) de la diagonal son 0.*

Por ejemplo, las siguientes matrices son triangulares

$$\begin{pmatrix} 1 & 0 & 0 \\ 5 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \quad \text{triangular inferior,}$$

$$\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \quad \text{triangular superior.}$$

Obsérvese que si A es una matriz de $m \times n$ como en (4.1), a los renglones de A se les pueden asociar

$$\begin{array}{rcl} R_1 & = & (a_{11}, \quad a_{12}, \quad \ldots, \quad a_{1n}) \\ R_2 & = & (a_{21}, \quad a_{22}, \quad \ldots, \quad a_{2n}) \\ \vdots & & \vdots \\ R_m & = & (a_{m1}, \quad a_{m2}, \quad \ldots, \quad a_{mn}), \end{array} \tag{4.2}$$

que son m vectores en $\mathbb{R}^n$, llamados vectores renglón; y a las columnas se les puede asociar

$$\begin{array}{rcl} C_1 & = & (a_{11}, \quad a_{21}, \quad \ldots, \quad a_{m1}) \\ C_2 & = & (a_{12}, \quad a_{22}, \quad \ldots, \quad a_{m2}) \\ \vdots & & \vdots \\ C_n & = & (a_{1n}, \quad a_{2n}, \quad \ldots, \quad a_{mn}), \end{array} \tag{4.3}$$

que son n vectores en $\mathbb{R}^m$, llamados vectores columna.

Si una matriz B se obtiene de otra A eliminando algunos renglones y/o columnas se dice que B es una submatriz de A.

4.2. El rango de una matriz

Definición 48 *Sea A la matriz descrita en* (4.1)*, se define el rango de A como la dimensión del subespacio vectorial en $\mathbb{R}^n$ generado por los vectores renglón, i.e, los descritos en* (4.2)*.*

Se probará, al final del capítulo, que esta dimensión es la misma que la del subespacio vectorial en $\mathbb{R}^m$ generado por los vectores columna. Obsérvese que si r es el rango, entonces $r \leq m$ (ya que son m vectores renglón) y también $r \leq n$ (ya que estos vectores viven en $\mathbb{R}^n$).

Por ejemplo, el rango de la siguiente matriz es 2

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 4 & 0 \end{pmatrix},$$

y de esta otra es 3

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 5 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

Operaciones elementales

A continuación se definen ciertas operaciones llamadas elementales para renglones (también se puede hacer lo mismo para columnas).

operación 1: **intercambio de renglones**.

operación 2: **multiplicación de un renglón por un escalar distinto de** 0.

operación 3: **sumar un renglón a otro.**

Definición 49 *Dos matrices A y B son equivalentes si se puede obtener una de otra mediante un número finito de operaciones elementales, se escribirá $A \sim B$.*

Obsérvese que $\sim$ es de equivalencia (ejercicio).

Ejemplo de operación 1:

$$\begin{pmatrix} 1 & 3 & 0 \\ 4 & \pi & 7 \end{pmatrix} \sim \begin{pmatrix} 4 & \pi & 7 \\ 1 & 3 & 0 \end{pmatrix},$$

de operación 2:

$$\begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 3 & 0 & 12 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix},$$

de operación 3:

$$\begin{pmatrix} 1 & 3 & 0 \\ 4 & \pi & 7 \end{pmatrix} \sim \begin{pmatrix} 5 & \pi+3 & 7 \\ 4 & \pi & 7 \end{pmatrix}.$$

Otro ejemplo

$$\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \\ z_1 & z_2 \end{pmatrix} \sim \begin{pmatrix} x_1 & x_2 \\ \alpha y_1 & \alpha y_2 \\ z_1 & z_2 \end{pmatrix} \sim \begin{pmatrix} x_1 + \alpha y_1 & x_2 + \alpha y_2 \\ y_1 & y_2 \\ \beta z_1 & \beta z_2 \end{pmatrix}$$

$$\sim \begin{pmatrix} x_1 + \alpha y_1 + \beta z_1 & x_2 + \alpha y_2 + \beta z_2 \\ y_1 & y_2 \\ z_1 & z_2 \end{pmatrix},$$

es decir, usando operaciones elementales, a un renglón se le puede sumar una combinación lineal de otros renglones.

Este procedimiento permite hacer ceros, por ejemplo:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -5 \\ 3 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -5 \\ 0 & -6 & -8 \end{pmatrix}$$

Obsérvese que siempre se puede hacer 0 debajo de una entrada distinta de 0 (si las entradas están en un campo), por ejemplo si $a_1 \neq 0$,

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \sim \begin{pmatrix} 1 & a_2/a_1 & a_3/a_1 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \sim \begin{pmatrix} 1 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

Teorema 4.2.1 *Las operaciones elementales no alteran el rango y por consiguiente 2 matrices equivalentes tienen el mismo rango.*

Demostración. Sean $R_1, R_2, \ldots, R_m$ los vectores renglón de una matriz A, entonces el rango de A resulta ser la dimensión del subespacio vectorial W, generado por estos vectores, *i.e.* $W =< R_1, R_2, \ldots, R_m >$. Hay que probar que cada una de las 3 operaciones no altera el rango:

La operación 1 ciertamente no lo cambia (ya que al intercambiar 2 renglones el espacio W no cambia).

Con respecto a la operación 2, si multiplicamos R_i por λ, se tiene $< R_1, R_2, \ldots, R_m >=< R_1, \ldots, \lambda R_i, \ldots, R_m >$, ya que los vectores que generan el primer subespacio están en el segundo, y viceversa,

$$R_i = \frac{1}{\lambda}(\lambda R_i).$$

Finalmente, en cuanto a la operación 3, si

$$W' =< R_1, \ldots, R_{j-1}, R_i + R_j, R_{j+1}, \ldots, R_m >,$$

ciertamente $W' \subset W$, ya que todas las combinaciones de los vectores que definen W', lo son también de $R_1, R_2, \ldots, R_m$. Finalmente, para probar que $W \subset W'$, basta probar que $R_j \in W'$, lo cual se sigue de la identidad

$$R_j = (R_i + R_j) - R_i.$$

□

Obsérvese que la prueba del teorema anterior muestra algo más general: *El subespacio vectorial generado por los vectores renglón no se altera bajo las operaciones elementales, i.e., es invariante.* Este teorema permite calcular el rango de una matriz al transformarla mediante operaciones elementales en otra, para la cual es fácil calcular el rango.

Definición 50 *Una matriz es escalonada si la primera entrada distinta de cero de cada renglón está más a la derecha de la primera entrada distinta de cero del renglón anterior.*

Ejemplos de matrices escalonadas:

$$\begin{pmatrix} 1 & 3 & \pi & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 9 & 4 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 5 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

Por otra parte, la matriz $\begin{pmatrix} 0 & 1 & 4 & 1 & 1 \\ 0 & 3 & 0 & 0 & 0 \end{pmatrix}$ no es escalonada.

Lema 4.2.2 *En una matriz escalonada, los vectores renglón distintos de* $\overline{0}$ *son linealmente independientes.*

Antes de probar el lema exhibimos un ejemplo:

$$\text{Sea} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

entonces, si

$$\lambda_1(1,1,1,1,1) + \lambda_2(0,1,1,1,1) + \lambda_3(0,0,0,1,1) = \overline{0},$$

se sigue que

$$\begin{cases} \lambda_1 & & & & & = & 0 \\ \lambda_1 & + & \lambda_2 & & & = & 0 \\ \lambda_1 & + & \lambda_2 & + & \lambda_3 & = & 0, \end{cases}$$

y $\lambda_1, \lambda_2, \lambda_3 = 0$.

Demostración del lema. En una matriz de $m \times n$ sean

$$1 \leq j_1 < j_2 < \cdots < j_k \leq n, \quad k \leq n,$$

tales que j_1 es el número de columna donde aparece el primer número $\neq 0$ del 1er renglón, j_2 el número de columna donde aparece el primer número $\neq 0$ del 2o renglón, etcétera.

Ahora si

$$\lambda_1 R_1 + \lambda_2 R_2 + \cdots + \lambda_m R_m = \overline{0},$$

donde R_i son los vectores renglón de la matriz. Considerando la ecuación escalar determinada por la coordenada j_1, se tiene $\lambda_1 = 0$. Con la definida por la coordenada j_2, se obtiene $\lambda_1 a_{1j_2} + \lambda_2 a_{2j_2} = 0$, por lo que $\lambda_2 = 0$, etcétera.

□

Por consiguiente, para calcular el rango de una matriz basta transformarla mediante operaciones elementales en un escalonada y contar los renglones de ésta distintos de 0.

Ejemplos

1. Sea
$$A = \begin{pmatrix} 1 & -1 & 3 & -5 \\ 2 & -3 & 4 & -10 \\ -3 & 3 & -9 & 15 \\ 3 & -3 & -6 & -4 \end{pmatrix},$$
haciendo ceros debajo de la entrada a_{11}, se tiene
$$A \sim \begin{pmatrix} 1 & -1 & 3 & -5 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -15 & 11 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 3 & -5 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & -15 & 11 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$
por lo que A tiene rango 3.

2. Sea A la matriz $\begin{pmatrix} 0 & 1 & 2 & 3 & 1 & 0 \\ 0 & -1 & -2 & -4 & 2 & 1 \\ 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 1 & 2 & 2 & 4 & 1 \end{pmatrix}$,
$$A \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & -1 & 3 & 1 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 6 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$
i.e, A tiene rango 3.

EJERCICIOS 4.2
1. Calcular el rango de las siguientes matrices

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & -1 \\ 3 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 2 & 0 \\ 2 & 3 & 1 & 4 \\ 1 & 5 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 1 & -1 & 3 \\ 1 & 2 & -1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 4 \\ 3 & 0 & 1 & 1 & -1 \end{pmatrix}.$$

2. Pruebe que la relación establecida en la Definición 49 es de equivalencia.

4.3. Permutaciones

Recordamos que una permutación de $I_n = \{1, 2, \ldots, n\}$ es una función biyectiva de I_n en I_n y que hay $n!$ permutaciones de I_n. El conjunto de ellas se denota por S_n, por ejemplo

$$S_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\},$$

$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}.$$

Escribiremos para una permutación $\sigma \in S_n$

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

o simplemente $\sigma = (\sigma(1), \sigma(2), \ldots, \sigma(n))$.

Permutaciones pares e impares

Consideramos primero el siguiente ejemplo: sea

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 1 & 5 & 6 & 2 & 4 \end{pmatrix},$$

diremos que $\sigma(2) = 7$ y $\sigma(3) = 1$, en el 2o renglón forman una inversión, ya que $2 < 3$, pero $\sigma(2) = 7 > \sigma(3) = 1$, también $\sigma(5)$ y $\sigma(7)$ lo hacen. Sin embargo, $\sigma(3)$ y $\sigma(6)$ no la forman, ya que $\sigma(3) = 1 < 2 = \sigma(6)$. En consecuencia, la definición formal general es la siguiente.

Definición 51 *Sea $\sigma \in S_n$, se dice que $\sigma(i)$ y $\sigma(j)$ forman inversión, si $i < j$ y $\sigma(i) > \sigma(j)$.*

Para contar el número de inversiones de una permutación dada basta contar cuantos números mayores a un número dado en el 2o renglón lo preceden. En el ejemplo anterior al 1 lo preceden 3 y 7, a 5 el 7, a 6 el 7, a 2, lo hacen 3, 7, 5 y 6, y a 4, 7, 5 y 6. Por lo que esta permutación tiene 11 inversiones.

Definición 52 *Se dice que una permutación es par si tiene un número par de inversiones, e impar si este número es impar (la paridad de la permutación es la propiedad de ser par o impar).*

Por ejemplo, la permutación $(1\ 2\ 3\ 4\ 5)$ es par, ya que tiene cero inversiones y la permutación $(2\ 1\ 3\ 4\ 5)$ es impar, ya que tiene una inversión $(\sigma(1) > \sigma(2))$. Nuestro ejemplo original $\sigma = (3\ 7\ 1\ 5\ 6\ 2\ 4)$ es impar, ya que tiene 11 inversiones.

Transposiciones

Definición 53 *Una transposición es una permutación que consiste en intercambiar 2 elementos.*

Por ejemplo

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 3 & 4 & 2 \end{pmatrix}.$$

Construiremos una nueva permutación a partir de una permutación dada, al intercambiar dos elementos en el 2o renglón, *i.e.*, si

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & r & \cdots & s & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(r) & \cdots & \sigma(s) & \cdots & \sigma(n) \end{pmatrix}$$

y τ es la transposición que intercambia $\sigma(r)$ y $\sigma(s)$, se obtiene una nueva permutación $\mu = \tau \cdot \sigma$

$$\mu = \begin{pmatrix} 1 & 2 & \cdots & r & \cdots & s & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(s) & \cdots & \sigma(r) & \cdots & \sigma(n) \end{pmatrix},$$

probaremos que si σ es par, entonces μ es impar, y que si σ es impar, μ es par, *i.e.*, σ y μ tienen paridad distinta.

Por ejemplo, si

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 1 & 5 & 6 & 2 & 4 \end{pmatrix} \quad \text{y} \quad \mu = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 7 & 5 & 1 & 6 & 2 & 4 \end{pmatrix},$$

se había mostrado que σ tiene 11 inversiones, entonces μ tiene 12, ya que tiene todas las de σ y una más, 1 con 5.

Lema 4.3.1 *Sea μ la permutación obtenida de σ mediante la transposición de 2 elementos consecutivos en el 2o renglón de σ, entonces σ y μ tienen distinta paridad.*

DEMOSTRACIÓN. Probamos que si σ tiene k inversiones, entonces μ tiene $k-1$ o $k+1$ inversiones. Sean

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & j & j+1 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(j) & \sigma(j+1) & \cdots & \sigma(n) \end{pmatrix}$$

y

$$\mu = \begin{pmatrix} 1 & 2 & \cdots & j & j+1 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(j+1) & \sigma(j) & \cdots & \sigma(n) \end{pmatrix}.$$

Todas las parejas que forman inversión en σ, también lo hacen en μ, salvo $\sigma(j)$ y $\sigma(j+1)$; éstas lo forman en σ si y sólo si no lo forman en μ.

□

Generalizamos ahora este resultado.

Teorema 4.3.2 *Sea μ la permutación obtenida de la permutación σ mediante la transposición de dos elementos en el 2o renglón de σ, entonces σ y μ tienen distinta paridad.*

DEMOSTRACIÓN. El caso en los números que se transponen son consecutivos se probó en el lema. Supongamos ahora que estos números están separados por s lugares, *i.e.*,

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & r & \cdots & r+s & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(r) & \cdots & \sigma(r+s) & \cdots & \sigma(n) \end{pmatrix}, \quad y$$

$$\mu = \begin{pmatrix} 1 & 2 & \cdots & r & \cdots & r+s & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(r+s) & \cdots & \sigma(r) & \cdots & \sigma(n) \end{pmatrix}.$$

En el 2o renglón de σ podemos mover $\sigma(r)$ a la derecha mediante s transposiciones de elementos consecutivos, hasta que ocupe el lugar $\sigma(r+s)$, obteniéndose

$$\sigma' = \begin{pmatrix} 1 & 2 & \cdots & r & \cdots & r+s & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(r+1) & \cdots & \sigma(r) & \cdots & \sigma(n) \end{pmatrix},$$

posteriormente mediante $s-1$ transposiciones de elementos consecutivos podemos mover $\sigma(r+s)$ a la izquierda, hasta que ocupe el lugar original de $\sigma(r)$, la permutación obtenida es precisamente μ. Por lo tanto se puede pasar de σ a μ mediante $2s-1$ transposiciones de elementos consecutivos, como la paridad cambia en cada paso y el número de éstos es impar, se sigue el resultado.

□

Teorema 4.3.3 *Sean $\sigma \in S_n$ y m el número de inversiones de σ, entonces σ se puede expresar como composición de m transposiciones.*

Demostración. La siguiente representación de σ ilustra la idea de la prueba

$$\sigma = \begin{pmatrix} \cdots\cdots\cdots \\ \cdots \curvearrowleft 1 \cdots \end{pmatrix}.$$

Se mueve el 1 a la izquierda en el 2o renglón con transposiciones de elementos consecutivos (todos estos números precedentes forman inversión con el 1), hasta que ocupe el 1er lugar. Obsérvese que este movimiento no altera las inversiones que forman otros números distintos de 1 entre sí. Posteriormente, se repite el proceso con el 2 hasta obtener $(1, 2, ...)$, etcétera. Se sigue que $\sigma_m \cdot \cdots \cdot \sigma_1 \cdot \sigma = I$, donde las σ_j denotan las transposiciones descritas. Nótese que el número de estos movimientos, *i.e.* m, es exactamente el número de inversiones de σ. Finalmente, $\sigma = \sigma_1 \cdot \cdots \cdot \sigma_m$.

□

Inversa de una permutación

Dada $\sigma \in S_n$, σ es una función biyectiva de I_n en I_n, por lo que existe $\sigma^{-1} : I_n \longrightarrow I_n$ y $\sigma^{-1} \in S_n$. Obsérvese que $\sigma(i) = k$ si y sólo si $\sigma^{-1}(k) = i$.

Si

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix},$$

la inversa se puede expresar como

$$\begin{pmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{pmatrix},$$

o como

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma^{-1}(1) & \sigma^{-1}(2) & \cdots & \sigma^{-1}(n) \end{pmatrix}.$$

Por ejemplo, si

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}, \quad \sigma^{-1} = \begin{pmatrix} 4 & 1 & 3 & 2 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}.$$

Obsérvese que las transposiciones son involuciones, *i.e.*, ellas son su propio inverso, ya que si τ es una transposición, entonces $\tau^2 = \tau \cdot \tau = Id$.

Podemos enumerar los inversos de los elementos de S_3.

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \sigma_1^{-1}, \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \sigma_2^{-1},$$

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \sigma_3^{-1},$$

$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_4^{-1} = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$\sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \sigma_5^{-1} = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$\sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \sigma_6^{-1},$$

por lo que $\sigma_4 = \sigma_5^{-1}$ y $\sigma_5 = \sigma_4^{-1}$.

Obsérvese que la asociación $\sigma \mapsto \sigma^{-1}$, es una biyección de S_n. Si σ, τ son distintas, entonces σ^{-1} y τ^{-1} lo son, y dada una permutación $\mu \in S_n$, μ es la imagen de su inversa bajo esta asociación.

Proposición 4.3.4 *Una permutación σ es par si y sólo si σ^{-1} lo es.*

Consideremos primero un ejemplo:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \cdot & 5 & \cdot & 1 & \cdot \end{pmatrix}, \quad \sigma^{-1} = \begin{pmatrix} \cdot & 5 & \cdot & 1 & \cdot \\ \cdot & 2 & \cdot & 4 & \cdot \end{pmatrix} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & 5 \\ 4 & \cdot & \cdot & \cdot & 2 \end{pmatrix},$$

i.e $\sigma(2) = 5$ y $\sigma(4) = 1$, forman inversión en σ si y sólo si $\sigma^{-1}(1) = 4$, lo hace con $\sigma^{-1}(5) = 2$, en σ^{-1}. La relación: el número *chico* en el primer renglón va al número *grande* en el segundo renglón se *preserva*.

Demostración de la proposición. La biyección $\sigma \mapsto \sigma^{-1}$ en S_n *preserva* inversiones:

$$i < j \quad \text{y} \quad r = \sigma(i) > \sigma(j) = s$$

$$\Longleftrightarrow$$

$$s < r \quad \text{y} \quad \sigma^{-1}(s) = j > i = \sigma^{-1}(r),$$

i.e., las parejas que forman inversión en σ se corresponden de manera biunívoca con las que lo hacen en σ^{-1}. Por consiguiente, el número de inversiones de σ es el mismo que el de σ^{-1}. □

Nótese que la inversión también se puede definir en términos del 1er renglón $i < j$ y $\sigma(i) > \sigma(j)$.

EJERCICIO 4.3
1. Encuentra la paridad de las siguientes permutaciones

$$\begin{pmatrix} 3 & 5 & 6 & 1 & 2 & 4 \end{pmatrix} \qquad \text{y} \qquad \begin{pmatrix} 1 & 7 & 8 & 2 & 4 & 3 & 5 & 6 \end{pmatrix}$$

4.4. Determinantes

El determinante es una función que le asocia a una matriz un número. Éstos son muy importantes en matemáticas, por ejemplo, en el estudio de las ecuaciones lineales. En muchos casos, estos números representan el área o el volumen de los lugares geométricos asociados a sus vectores renglón, es decir, los correspondientes paralelogramos o paralelepípedos, como se muestra al final de este capítulo.

Definición 54 *Se define el determinante de una matriz cuadrada*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

$$\textit{como} \qquad \det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

donde

$$\varepsilon(\sigma) = \begin{cases} 1 & \textit{si} \quad \sigma \textit{ es par}, \\ -1 & \textit{si} \quad \sigma \textit{ es impar}. \end{cases}$$

En particular, para el caso $n = 2$, esta definición es más simple y familiar.

Definición 55 *Sea*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matriz con entradas reales, al número $ad - bc$ *se le llama determinante de* A *y se le denota por* $\det A$ *o por* $|A|$.

Ejemplos

1. Si

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

entonces $det\, A = a_{11}a_{22} - a_{12}a_{21}$, ya que las permutaciones de $\{1, 2\}$ son

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{que es par, y} \quad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad \text{que es impar.}$$

2. Calculamos el determinante de la matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Calculamos primero las permutaciones de $\{1, 2, 3\}$, y sus paridades.

$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ es par (cero inversiones), genera el sumando $a_{11}a_{22}a_{33}$,

$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ es impar (una inversión), genera $-a_{11}a_{23}a_{32}$,

$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ es impar (una inversión), genera $-a_{12}a_{21}a_{33}$,

$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ es par (dos inversiones), genera $a_{12}a_{23}a_{31}$,

$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ es par (dos inversiones), genera $a_{13}a_{21}a_{32}$, y finalmente

$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ es impar (tres inversiones), define el sumando $-a_{13}a_{22}a_{31}$.

Por consiguiente,

$$det\, A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

Volviendo al caso general de matrices de $n \times n$, obsérvese que los sumandos están hechos por factores que son uno por cada renglón (ya que los primeros subíndices son $1, 2, \ldots, n$), y uno por cada columna. Esto se sigue ya que los segundos subíndices son $\sigma(1), \sigma(2), \ldots, \sigma(n)$, y σ es una permutación.

Dada una matriz diagonal A, su determinante es $a_{11}\, a_{22} \ldots a_{nn}$, *i.e.*, el producto de los elementos de la diagonal. Esto se sigue, ya que todos los sumandos donde aparecen factores $a_{j\sigma(j)}$ con $\sigma(j) \neq j$ son 0, esto es, la única permutación que produce sumandos distintos de 0 es la identidad.

Proposición 4.4.1 *Sea A una matriz triangular $A = (a_{ij})$, $1 \leq i, j \leq n$, entonces $det\, A = a_{11}a_{22} \cdots a_{nn}$.*

Demostración. Probamos la proposición para triangulares superiores (elementos debajo de la diagonal son 0), el otro caso es análogo. Se afirma que los sumandos donde aparece $a_{j\sigma(j)}$, $\sigma(j) \neq j$ son todos 0. Para probar la afirmación consideramos dos casos, donde σ es fija y no es la identidad.

Caso 1: $j > \sigma(j)$. En este caso por ser A triangular superior $a_{j\sigma(j)} = 0$, y se sigue la afirmación.

Caso 2: $j < \sigma(j)$. En este caso, alguna otra i, de los primeros n naturales, cumple $i > \sigma(i)$. De otra manera, $\sigma(n) \geq n$, y $\sigma(n) = n$. Asimismo $\sigma(n-1) \geq n-1$, por lo cual $\sigma(n-1) = n-1$, procediendo de esta manera, se tiene $\sigma(j) = j$, lo cual contradice la hipótesis $j < \sigma(j)$. Se concluye que existe dicha i, y entonces $a_{i\sigma(i)} = 0$ (por ser triangular superior).

□

4.5. Propiedades de los determinantes

Propiedad 1

Si el vector renglón R_i de una matriz A es la suma de 2 vectores renglón, digamos $R_i = R_i^{'} + R_i^{''}$, entonces

$$|A| = |A^{'}| + |A^{''}|,$$

donde A' *y* A'' *son las matrices cuyos renglones i-ésimos están determinados por* R_i' *y* R_i'', *respectivamente y los demás renglones son los mismos que los de* A.

Ejemplo

$$\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = (a+a')d-(b+b')c = ad-bc+a'd-b'c = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c' & d' \end{vmatrix}.$$

DEMOSTRACIÓN DE LA PROPIEDAD 1. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{4.4}$$

Si el vector renglón R_i de A es $R_i' + R_i''$, podemos escribir

$$(a_{i1}, a_{i2}, \ldots, a_{in}) = (a_{i1}', a_{i2}', \ldots, a_{in}') + (a_{i1}'', a_{i2}'', \ldots, a_{in}''),$$

y se tiene que

$$\begin{aligned} |A| &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots (a_{i\sigma(i)}' + a_{i\sigma(i)}'') \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)}' \cdots a_{n\sigma(n)} + \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)}'' \cdots a_{n\sigma(n)} \\ &= |A'| + |A''|. \end{aligned}$$

□

Esta propiedad nos permite simplificar determinantes, por ejemplo:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+0 & b_1+0 & c_1+0 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & 0 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} 0 & 0 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & 0 & 0 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} 0 & b_1 & 0 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} 0 & 0 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Propiedad 2

Si el vector renglón R_i *de la matriz* A *es de la forma* $R_i = \lambda R_i^{'}$, *entonces*

$$|A| = \lambda |A^{'}|,$$

donde $A^{'}$ *es la matriz que se obtiene de* A *cambiando el renglón definido por* R_i *por el renglón asociado a* $R_i^{'}$.

Ejemplos

1. $\begin{vmatrix} \lambda a & \lambda b \\ c & d \end{vmatrix} = \lambda ad - \lambda bc = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

2. Observando que en la expresión del determinante aparece solamente un factor de la forma a_{1j}, por cada sumando, $1 \leq j \leq 3$, se tiene

$$\begin{vmatrix} \lambda a_1 & \lambda b_1 & \lambda c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \lambda \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Demostración de la Propiedad 2. Sea A como en (4.4) y $R_i = \lambda R_i^{'}$, donde

$$R_i = (a_{i1}, a_{i2}, \ldots, a_{in}) = \lambda(a_{i1}^{'}, a_{i2}^{'}, \ldots, a_{in}^{'}),$$

entonces

$$\begin{aligned} |A| &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots \lambda a_{i\sigma(i)}^{'} \cdots a_{n\sigma(n)} \\ &= \lambda \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)}^{'} \cdots a_{n\sigma(n)} \\ &= \lambda |A^{'}|. \end{aligned}$$

□

Propiedad 3

Si la matriz $A^{'}$ *se obtiene intercambiando dos renglones de otra matriz* A, *entonces*

$$|A^{'}| = -|A|.$$

Ejemplo

$$\text{Sean} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{y} \quad A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix},$$

entonces $|A'| = cb - da = -(ad - bc) = -|A|$.

DEMOSTRACIÓN DE LA PROPIEDAD 3. Sean

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad \text{y} \quad A' = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix},$$

donde

$$\begin{array}{rcll} b_{ij} & = & a_{ij} & \forall\, i, j, \quad i \neq r, \quad i \neq s, \\ b_{rj} & = & a_{sj} & \forall\, j, \\ b_{sj} & = & a_{rj} & \forall\, j, \end{array}$$

i.e. A y A' son la misma matriz excepto por los renglones r y s intercambiados.

Ahora dada $\sigma \in S_n$, se define $\tau \in S_n$ como sigue:

$$\begin{array}{rcll} \tau(i) & = & \sigma(i), & \forall\, i \neq r, s, \\ \tau(r) & = & \sigma(s), & \\ \tau(s) & = & \sigma(r), & \end{array}$$

es decir,

$$\tau = \begin{pmatrix} 1 & \cdots & r & \cdots & s & \cdots & n \\ \sigma(1) & \cdots & \sigma(s) & \cdots & \sigma(r) & \cdots & \sigma(n) \end{pmatrix}.$$

Nótese que τ tiene distinta paridad de σ, ya que se obtiene de ésta mediante una transposición.

Finalmente

$$\begin{aligned} |A| &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{r\sigma(r)} \cdots a_{s\sigma(s)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{1\sigma(1)} \cdots b_{s\sigma(r)} \cdots b_{r\sigma(s)} \cdots b_{n\sigma(n)} \\ &= \sum_{\tau \in S_n} -\varepsilon(\sigma) b_{1\tau(1)} \cdots b_{s\tau(s)} \cdots b_{r\,\tau(r)} \cdots a_{n\tau(n)} \\ &= -|A'|. \end{aligned}$$

Si σ recorre todas las permutaciones, τ también, ya que $\sigma \mapsto \tau$ es una biyección en S_n.

□

Propiedad 4

Sea $A = (a_{ij})$ la matriz identidad, esto es, $a_{ij} = \delta_{ij}$, donde

$$\delta_{ij} = \begin{cases} 1 & si \quad i = j, \\ 0 & si \quad i \neq j. \end{cases}$$

Entonces, $|A| = 1$, a la función δ_{ij} se le llama la delta de Kronecker.

Demostración. Cada sumando $a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ es 0 salvo si $\sigma(j) = j \ \ \forall j$, *i.e.*, si $\sigma = Id$.

□

Ejemplo

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

Propiedad 5

Si una matriz cuadrada tiene 2 renglones iguales su determinante es 0.

Demostración. Sea A dicha matriz y A' la que se obtiene intercambiando los 2 renglones que son iguales *i.e.*, $A' = A$, se sigue entonces de la Propiedad 3 que $|A| = -|A'| = -|A|$, por lo que $2|A| = 0$ y $|A| = 0$.

□

Propiedad 6

Si una matriz cuadrada tiene un renglón de ceros, entonces su determinante es cero.

Demostración. En la definición de determinante aparece en cada sumando un factor de cada uno de los renglones, *i.e.*, cada sumando tiene un factor 0, *i.e.*, todos lo sumandos son 0.

Otra demostración se obtiene usando la Propiedad 2: se puede factorizar 0 del renglón de ceros y se tiene $|A| = 0 \cdot |A|$, *i.e.*, $|A| = 0$.

□

Ejemplo

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 + \lambda a_3 & b_2 + \lambda b_3 & c_2 + \lambda c_3 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & c_1 \\ \lambda a_3 & \lambda b_3 & \lambda c_3 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Las igualdades son consecuencia de las Propiedades 1, 2 y 5. La siguiente propiedad es una generalización de este ejemplo.

Propiedad 7

Si la matriz A' *se obtiene de la matriz cuadrada* A *sumando a un renglón un multiplo de otro, entonces* $|A'| = |A|$.

Demostración. Supóngase que al renglón R_i se le sumó el renglón R_k multiplicado por λ, entonces por las Propiedad 1 y la Propiedad 2 se tiene que

$$|A'| = |A| + \lambda|B|,$$

donde B es una matriz igual a A excepto por el renglón i que es el mismo que el renglón k. Finalmente la Propiedad 5 establece que $|B| = 0$.

□

Iterando este proceso se tiene

Corolario 4.5.1 *Si a un renglón de una matriz cuadrada* A *le sumamos una combinación lineal de los demás renglones, obteniendo una matriz* A'*, entonces* $|A'| = |A|$.

Corolario 4.5.2 *Si los vectores renglón de una matriz cuadrada* A *son linealmente dependientes, entonces* $|A| = 0$.

Demostración. Las hipótesis implican que alguno de los vectores renglón es combinación lineal de los otros, restando a dicho renglón esa combinación lineal se obtiene un renglón de ceros.

□

Por ejemplo, si

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 6 & 9 & 12 \end{pmatrix},$$

entonces $(6, 9, 12) = 2(1, 2, 3) + (4, 5, 6)$ y $|A| = 0$, ya que

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 6 - 2 \cdot 1 - 4 & 9 - 2 \cdot 2 - 5 & 12 - 2 \cdot 3 - 6 \end{vmatrix}.$$

Definición 56 *Dada una matriz* $A = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq m$, *se define otra matriz llamada la transpuesta de* A, *denotada por* ${}^tA$ *(o* A^t*), como sigue*

$$ {}^tA = (b_{ij}), \quad \textit{donde} \quad b_{ij} = a_{ji}, $$

Por ejemplo, si

$$ A = \begin{pmatrix} a & 2 & 0 \\ \pi & \epsilon & 1 \end{pmatrix}, \quad \text{entonces} \quad {}^tA = \begin{pmatrix} a & \pi \\ 2 & \epsilon \\ o & 1 \end{pmatrix}. $$

Es decir, el 1er renglón se transforma en la 1a columna y el 2o renglón en la 2a columna.

Propiedad 8

$$ |A| = |{}^tA|. $$

Ejemplo

$$ \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc = ad - cb = \begin{vmatrix} a & c \\ b & d \end{vmatrix} $$

Demostración de la Propiedad 8. Sean $A = (a_{ij})$ y ${}^tA = (b_{ij})$, *i.e.*, $a_{ij} = b_{ji}$ $\forall\, i, j$. Ahora

$$ |A| = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n}, $$

como $b_{\sigma(j)j} = b_{i\sigma^{-1}(i)}$, donde $\sigma(j) = i$; se pueden reordenar los factores de cada sumando y se tiene

$$ b_{\sigma(1)1} b_{\sigma(2)2} \cdots b_{\sigma(n)n} = b_{1\sigma^{-1}(1)} \cdots b_{n\sigma^{-1}(n)}. $$

Además $S_n = \{\sigma_1, \ldots, \sigma_{n!}\} = \{\sigma_1^{-1}, \ldots, \sigma_{n!}^{-1}\}$ y $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$, en virtud de la Proposición 4.3.4, por lo que

$$ |A| = \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)} = |{}^tA|. $$

□

Corolario 4.5.3 *Las propiedades* 1 *a* 7 *para renglones son también válidas para columnas.*

Ejemplo

$$\begin{vmatrix} a_1 & \lambda b_1 & c_1 \\ a_2 & \lambda b_2 & c_2 \\ a_3 & \lambda b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ \lambda b_1 & \lambda b_2 & \lambda b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \lambda \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \lambda \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

4.6. Desarrollo por menores

Este método nos permite expresar el determinante de una matriz de $n \times n$ como la suma de determinantes de matrices de $(n-1) \times (n-1)$.

Dada una matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

se denota por A_{ij} la matriz obtenida de A eliminando el renglón i y la columna j.

Por ejemplo, si

$$A = \begin{pmatrix} a & \pi & 0 \\ 2 & 4 & 7 \\ 9 & 6 & 8 \end{pmatrix}, \tag{4.5}$$

entonces

$$A_{11} = \begin{pmatrix} 4 & 7 \\ 6 & 8 \end{pmatrix}, \; A_{23} = \begin{pmatrix} a & \pi \\ 9 & 6 \end{pmatrix}, \; A_{32} = \begin{pmatrix} a & 0 \\ 2 & 7 \end{pmatrix},$$

etcétera.

Definición 57 *Al determinante* $|A_{ij}|$ *de una matriz cuadrada* A *se le llama el menor del elemento* a_{ij}.

Por ejemplo, si A es la matriz (4.5), se tiene $|A_{11}| = -10$ y $|A_{32}| = 7a$.

Podemos desarrollar el determinante de una matriz de 3×3 con respecto a los menores del 1er renglón. Sea

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix},$$

entonces

$$\begin{aligned} |A| &= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}\\ &= a_{11}|A_{11}|-a_{12}|A_{12}|+a_{13}|A_{13}|. \end{aligned}$$

Probaremos que este proceso se puede generalizar para cualquier renglón o columna.

Lema 4.6.1 *Sea*

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix},$$

tal que $a_{nj}=0 \quad \forall\, j\neq n$, *entonces*

$$|A|=a_{nn}|A_{nn}|.$$

Ejemplo

$$\begin{vmatrix} 1 & 0 & 2 & 4\\ 3 & 7 & 1 & 8\\ -1 & 4 & 0 & 2\\ 0 & 0 & 0 & 7\end{vmatrix} = 7\begin{vmatrix} 1 & 0 & 2\\ 3 & 7 & 1\\ -1 & 4 & 0\end{vmatrix}.$$

Demostración.

$$|A|=\sum_{\sigma\in S_n}\varepsilon(\sigma)a_{1\sigma(1)}\cdots a_{n\sigma(n),}$$

como $a_{n\sigma(n)}=0$ si $\sigma(n)\neq n$, se tiene

$$|A|=\sum_{\sigma\in S_{n-1}}\varepsilon(\sigma)a_{1\sigma(1)}\cdots a_{n-1\sigma(n-1)}a_{nn}=a_{nn}|A_{nn}|,$$

ya que las permutaciones que fijan a n están en correspondencia biunívoca con las permutaciones de S_{n-1} (el número de inversiones también es el mismo).

□

Corolario 4.6.2 *Sea $A = (a_{ij})$ una matriz de $n \times n$, tal que $a_{jn} = 0$ para los naturales j menores a n, entonces $|A| = a_{nn}|A_{nn}|$.*

DEMOSTRACIÓN. Aplicando el Lema 4.6.1 a A^t, se tiene $|A^t| = a_{nn}|(A^t)_{nn}|$. Ahora, en virtud de la Propiedad 8 se sigue que $|A| = a_{nn}|A_{nn}|$. En este último paso usamos el hecho de que $(A_{nn})^t = (A^t)_{nn}$ (ejercicio).

□

Ejemplo

$$\text{Si} \quad A = \begin{pmatrix} 2 & 3 & 1 \\ 5 & 7 & 2 \\ 0 & 2 & 7 \end{pmatrix}, \quad \text{entonces} \quad A^t = \begin{pmatrix} 2 & 5 & 0 \\ 3 & 7 & 2 \\ 1 & 2 & 7 \end{pmatrix}.$$

Por lo cual

$$(A_{33})^t = \begin{pmatrix} 2 & 5 \\ 3 & 7 \end{pmatrix} = (A^t)_{33}.$$

Lema 4.6.3 *Si en una matriz cuadrada todos los elementos distintos de a_{ij} en el renglón i son cero, entonces*

$$|A| = (-1)^{i+j} a_{ij} |A_{ij}|.$$

DEMOSTRACIÓN. Intercambiamos renglones y columnas consecutivas hasta obtener la situación del Lema 4.6.1. Específicamente sean

$$A = \left(\begin{array}{ccc|c|ccc} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & & & & & \vdots \\ a_{i-1,1} & & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & & a_{i-1.n} \\ \hline a_{i,1} & & a_{i,j-1} & a_{i,j} & a_{i,j+1} & & a_{i,n} \\ \hline a_{i+1,1} & & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & & a_{i+1,n} \\ \vdots & & & & & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{n,n} \end{array}\right) \quad \text{y}$$

$$A' = \left(\begin{array}{ccccccc|c} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} & a_{1,j} \\ \vdots & & & & & \vdots & \vdots \\ a_{i-1,1} & & a_{i-1,j-1} & a_{i-1,j+1} & & a_{i-1,n} & a_{i-1,j} \\ a_{i+1,1} & & a_{i+1,j-1} & a_{i+1,j+1} & & a_{i+1,n} & a_{i+1,j} \\ \vdots & & & & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} & a_{n,j} \\ \hline a_{i,1} & & a_{i,j-1} & a_{i,j+1} & & a_{i,n} & a_{i,j} \end{array}\right).$$

Obsérvese que se pueden intercambiar renglones consecutivos, de tal manera que el renglón R_i de A se transforme en el renglón n-ésimo, esto se obtiene mediante $n-i$ intercambios. Análogamente se puede intercambiar la columna C_j por la columna C_{j+1}, después por C_{j+2} hasta que ocupe la columna C_j el lugar de C_n. Es claro entonces que A' se obtiene de A mediante $n-i$ intercambios de renglones y $n-j$ de columnas. Finalmente, como cada intercambio de columnas o reglones cambia el signo del determinante, se tiene por el Lema 4.6.1 que

$$|A| = (-1)^{2n-i-j}|A'| = (-1)^{i+j}|A'| = (-1)^{i+j}a_{ij}|A'_{nn}| = (-1)^{i+j}a_{ij}|A_{ij}|,$$

puesto que $(-1)^k = (-1)^{-k} = \frac{1}{(-1)^k}$, y evidentemente la submatriz A'_{nn} es igual a la submatriz A_{ij} (ambas consisten de eliminar el renglón i y la columna j en A).

□

Corolario 4.6.4 *Si en una matriz cuadrada todos los elementos distintos de a_{ij} en la columna j son cero, entonces*

$$|A| = (-1)^{i+j}a_{ij}|A_{ij}|.$$

Demostración. Exactamente la misma prueba del Lema 4.6.3, aplicando el Corolario 4.6.2 en lugar de Lema 4.6.1, implica el resultado.

□

Estos resultados permiten desarrollar el determinante de una matriz utilizando cualquier renglón o columna.

Propiedad 9

Sea A una matriz de $n \times n$, y $1 \leq i, j \leq n$, entonces

$$\begin{aligned} |A| &= (-1)^{i+1}a_{i,1}|A_{i,1}| + (-1)^{i+2}a_{i,2}|A_{i,2}| + \cdots + (-1)^{i+n}a_{i,n}|A_{i,n}| \\ &= (-1)^{1+j}a_{1,j}|A_{1,j}| + (-1)^{2+j}a_{2,j}|A_{2,j}| + \cdots + (-1)^{n+j}a_{n,j}|A_{n,j}|. \end{aligned}$$

Antes de probar esta propiedad damos algunos ejemplos. Calculamos el siguiente determinante desarrollando con respecto a la 2a columna:

$$\begin{vmatrix} 5 & 7 & 9 \\ 2 & 1 & 0 \\ 3 & 4 & 5 \end{vmatrix} = (-1)^{1+2} \cdot 7 \begin{vmatrix} 2 & 0 \\ 3 & 5 \end{vmatrix} + (-1)^{2+2} \begin{vmatrix} 5 & 9 \\ 3 & 5 \end{vmatrix} + (-1)^{3+2} \cdot 4 \begin{vmatrix} 5 & 9 \\ 2 & 0 \end{vmatrix}$$

$$= -70 - 2 - 4(-18) = 0.$$

En la siguiente matriz desarrollamos con respecto al 3er renglón:

$$\begin{vmatrix} 1 & 0 & 2 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 2 & 1 & 0 & 3 \end{vmatrix} = (-1)^{3+3} \cdot 2 \begin{vmatrix} 1 & 0 & 0 \\ 3 & 0 & 0 \\ 2 & 1 & 3 \end{vmatrix} + (-1)^{3+4} \begin{vmatrix} 1 & 0 & 2 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{vmatrix}$$

$$= 2 \cdot 0 - [1(-1) + 2(3)] = -5,$$

DEMOSTRACIÓN DE LA PROPIEDAD 9. Probamos primero para el caso de los renglones, sea $R_i = (a_{i_1}, a_{i_2}, \ldots, a_{i_n})$ el vector renglón i-ésimo de A, entonces si los renglones se denotan también por $R_1, R_2, \ldots, R_n$ (abusando de la notación), se tiene por la Propiedad 1 que

$$|A| = \begin{vmatrix} R_1 \\ R_2 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{vmatrix} = \begin{vmatrix} R_1 \\ R_2 \\ \vdots \\ a_{i_1}\ 0\ \cdots\ 0 \\ \vdots \\ R_n \end{vmatrix} + \begin{vmatrix} R_1 \\ R_2 \\ \vdots \\ 0\ a_{i_2}\ 0\ \cdots\ 0 \\ \vdots \\ R_n \end{vmatrix} + \cdots + \begin{vmatrix} R_1 \\ R_2 \\ \vdots \\ 0\ \cdots\ 0\ a_{i_n} \\ \vdots \\ R_n \end{vmatrix}$$

$$= \sum_{k=1}^{n} (-1)^{i+k} a_{i_k} |A_{i_k}|,$$

aplicando el Lema 4.6.3. El mismo argumento se aplica para columnas, usando ahora el Corolario 4.6.4:

$$|A| = \begin{vmatrix} & a_{i_j} & \\ & 0 & \\ C_1\ C_2 \cdots & 0 & \cdots C_n \\ & \vdots & \\ & 0 & \end{vmatrix} + \cdots + \begin{vmatrix} & 0 & \\ & 0 & \\ C_1\ C_2 \cdots & \vdots & \cdots C_n \\ & 0 & \\ & a_{n_j} & \end{vmatrix},$$

donde C_k denota la columna k-ésima, etcétera. □

EJERCICIOS 4.6

1. Calcular el determinante de las siguientes matrices

$$\begin{pmatrix} 1 & 1 & 2 & -1 \\ 3 & 0 & 1 & -1 \\ -1 & 2 & 3 & 1 \\ 1 & 2 & -1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 & -1 \\ 1 & 3 & 2 & 1 \\ -1 & 1 & 0 & 3 \\ 2 & -1 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & -1 & 2 & 1 \\ 0 & 2 & 0 & 1 & 2 \\ 1 & 3 & 2 & 1 & 1 \\ 0 & 1 & -2 & 0 & 2 \\ 1 & 3 & 1 & 3 & 0 \end{pmatrix}.$$

2. Terminar la prueba del Corolario 4.6.2.

4.7. Cálculo de determinantes

Calcular un determinante a partir de la definición no es muy apropiado, por ejemplo el determinante de una matriz de 7×7 involucra $7! = 5040$ sumandos. El método de menores simplifica un poco esta situación, pero no mucho, por ejemplo, encontrar un determinante de 8×8 involucra sumar 8 determinantes de 7×7, etcétera.

En general, la manera más eficaz es simplificar la matriz haciéndola triangular usando las propiedades 1 a 8. En algunos casos es conveniente usar también el método de menores.

Ejemplos

1. Primero, se hacen ceros *debajo* de a_{11}

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 2 & -1 & 2 & -1 \\ 1 & 2 & -1 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 \\ 0 & -3 & 0 & -3 \\ 0 & 1 & -2 & -3 \end{vmatrix} = 6 \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -2 & -3 \end{vmatrix},$$

$$= 6 \left[- \begin{vmatrix} 1 & 1 \\ 1 & -3 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 1 & -2 \end{vmatrix} \right] = 6[4 - 2] = 12.$$

En el último paso, se desarrolló por menores en el primer renglón.

2. Haciendo ceros

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 5 & 9 \\ 0 & 3 & 9 & 19 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 5 & 9 \\ 3 & 9 & 19 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \\ 0 & 3 & 10 \end{vmatrix}.$$

$$= 10 - 9 = 1.$$

En el último paso, se desarrolló por menores en la 1a columna.

3. Sumando las 4 columnas

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 10 & 2 & 3 & 4 \\ 10 & 3 & 4 & 1 \\ 10 & 4 & 1 & 2 \\ 10 & 1 & 2 & 3 \end{vmatrix} = 10 \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 1 \\ 1 & 4 & 1 & 2 \\ 1 & 1 & 2 & 3 \end{vmatrix}$$

$$= 10 \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & -3 \\ 0 & 2 & -2 & -2 \\ 0 & -1 & -1 & -1 \end{vmatrix} = 20 \begin{vmatrix} 1 & 1 & -3 \\ 1 & -1 & -1 \\ -1 & -1 & -1 \end{vmatrix} = 20 \begin{vmatrix} 1 & 1 & -3 \\ 0 & -2 & 2 \\ 0 & 0 & -4 \end{vmatrix} = 160.$$

Determinantes tipo Van der Monde

Multiplicando por a_1 la 2a columna y restándola de la 3a y posteriormente multiplicando la 1a columna por a_1 y restándola de la 2a, se tiene

$$\begin{vmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{vmatrix} = \begin{vmatrix} 1 & a_1 & 0 \\ 1 & a_2 & a_2^2 - a_1 a_2 \\ 1 & a_3 & a_3^2 - a_1 a_3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_1 a_2 \\ 1 & a_3 - a_1 & a_3^2 - a_1 a_3 \end{vmatrix}$$

$$= (a_2 - a_1)(a_3 - a_1) \begin{vmatrix} 1 & a_2 \\ 1 & a_3 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2).$$

$$\begin{vmatrix} 1 & a_1 & a_1^2 & a_1^3 \\ 1 & a_2 & a_2^2 & a_2^3 \\ 1 & a_3 & a_3^2 & a_3^3 \\ 1 & a_4 & a_4^2 & a_4^3 \end{vmatrix} = \begin{vmatrix} 1 & a_1 & a_1^2 & 0 \\ 1 & a_2 & a_2^2 & a_2^3 - a_1 a_2^2 \\ 1 & a_3 & a_3^2 & a_3^3 - a_1 a_3^2 \\ 1 & a_4 & a_4^2 & a_4^3 - a_1 a_4^2 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & a_1 & 0 & 0 \\ 1 & a_2 & a_2^2 - a_1a_2 & a_2^3 - a_1a_2^2 \\ 1 & a_3 & a_3^2 - a_1a_3 & a_3^3 - a_1a_3^2 \\ 1 & a_4 & a_4^2 - a_1a_4 & a_4^3 - a_1a_4^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_1a_2 & a_2^3 - a_1a_2^2 \\ 1 & a_3 - a_1 & a_3^2 - a_1a_3 & a_3^3 - a_1a_3^2 \\ 1 & a_4 - a_1 & a_4^2 - a_1a_4 & a_4^3 - a_1a_4^2 \end{vmatrix}$$

$$= (a_2 - a_1)(a_3 - a_1)(a_4 - a_1) \begin{vmatrix} 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \\ 1 & a_4 & a_4^2 \end{vmatrix} = \prod_{j<k} (a_k - a_j),$$

donde $\prod$ significa producto.

EJERCICIOS 4.7

1. Generalizar el último ejemplo probando que

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots, & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{j<k} (a_k - a_j).$$

2. Calcular los determinantes de las siguientes matrices

$$\begin{pmatrix} 1 & -1 & 2 & 3 & 4 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & -1 & 0 & 1 & 4 \\ 5 & 1 & 0 & 2 & 7 \\ 1 & 2 & 1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & 0 & 2 & 4 \\ 1 & -1 & 1 & 1 & 1 \\ 4 & 3 & 2 & 1 & 0 \\ 5 & 1 & -1 & 0 & 7 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

4.8. Caracterización del rango de una matriz usando determinantes

Probaremos que el rango de una matriz es también la dimensión del subespacio vectorial generado por los vectores columna.

Teorema 4.8.1 *Un conjunto de vectores* $\{A_1, A_2, \ldots, A_k\}$ *en* $\mathbb{R}^n$ *son linealmente dependientes* $\Leftrightarrow$ *todos los determinantes de las submatrices de* $k \times k$ *obtenidos al fijar* k *coordenadas de los vectores* $A_1, A_2, \ldots, A_k$, *son cero.*

Exhibimos primero unos ejemplos que ilustran las técnicas de la prueba. Primero, si $A = (a_1, a_2, \ldots, a_n)$ y $B = (b_1, b_2, \ldots, b_n)$ son vectores linealmente dependientes, entonces existen escalares $\alpha, \beta \in \mathbb{R}$ ∋ $\alpha A + \beta B = 0$ y alguno de ellos es distinto de 0, digamos $\beta \neq 0$. Bajo estas hipótesis, se tiene entonces que $\alpha a_i + \beta b_i = 0 \; \forall \; i \in \{1, \ldots, n\}$, y

$$0 = \begin{vmatrix} a_i & a_j \\ 0 & 0 \end{vmatrix} = \begin{vmatrix} a_i & a_j \\ \alpha a_i + \beta b_i & \alpha a_j + \beta b_j \end{vmatrix} = \beta \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix} \; \therefore \; \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix} = 0 \; \forall i, j.$$

Por otra parte, sean $A = (a_1, \ldots, a_n)$ y $B = (b_1, \ldots, b_n)$ vectores que satisfacen la siguiente condición

$$\begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix} = 0 \quad \forall \, i, j, \; i \neq j.$$

Si $A = 0$, A, B son linealmente dependientes, por lo que podemos suponer $A \neq 0$, digamos $a_j \neq 0$. Se tiene entonces que

$$-b_j A + a_j B = (-b_j a_1 + a_j b_1, \ldots, -b_j a_n + a_j b_n) = \overline{0},$$

como $a_j \neq 0$, se sigue que A y B son linealmente dependientes (claramente, la condición $a_i b_i - a_i b_i = 0$ siempre se cumple).

Demostración del Teorema 4.8.1. Probamos primero la necesidad, sean

$$\begin{aligned} A_1 &= (a_{11}, a_{12}, \ldots, a_{1n}), \\ A_2 &= (a_{21}, a_{22}, \ldots, a_{2n}), \\ &\vdots \\ A_k &= (a_{k1}, a_{k2}, \ldots, a_{kn}), \end{aligned}$$

vectores linealmente dependientes, entonces existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$, alguno distinto de 0, tales que

$$\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_k A_k = \overline{0}.$$

Sin perder generalidad, $\alpha_k \neq 0$. Tomando coordenadas $j = 1, 2, \ldots, n$, se tiene que $\alpha_1 a_{1,j} + \alpha_2 a_{2,j} + \cdots \alpha_k a_{k,j} = 0$.

Finalmente, sean $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, k coordenadas distintas, entonces

$$0 = \begin{vmatrix} a_{1,i_1} & \cdots & a_{1,i_k} \\ a_{2,i_1} & & a_{2,i_k} \\ \vdots & & \vdots \\ \alpha_1 a_{1,i_1} + \alpha_2 a_{2,i_1} + \cdots + \alpha_k a_{k,i_1} & \cdots & \alpha_1 a_{1,i_k} + \alpha_2 a_{2,i_k} + \cdots + \alpha_k a_{k,i_k} \end{vmatrix}$$

$$= \alpha_k \begin{vmatrix} a_{1,i_1} & \cdots & a_{1,i_k} \\ a_{2,i_1} & \cdots & a_{2,i_k} \\ \vdots & & \vdots \\ a_{k,i_1} & \cdots & a_{k,i_k} \end{vmatrix},$$

y se sigue que cualquier submatriz de $k \times k$ tiene determinante igual a 0, ya que $\alpha_k \neq 0$.

Probamos ahora el recíproco por inducción. Si $k = 1$ el resultado es trivial, un vector cuyas coordenadas son cero, es el vector $\overline{0}$ que es linealmente dependiente.

Suponemos cierto el teorema para $k-1$ vectores y probamos para k vectores, digamos $A_1, \ldots, A_{k-1}, A_k$. Si todas las submatrices de $k-1 \times k-1$ obtenidas de los vectores $A_1, \ldots, A_{k-1}$ tienen determinante 0, se sigue de la hipótesis de inducción que $A_1, \ldots, A_{k-1}$ son linealmente dependientes, y por lo tanto $A_1, \ldots, A_{k-1}, A_k$ también lo son. De otra manera, existe una submatriz con determinante distinto de 0 de la forma

$$\begin{vmatrix} a_{1,i_1} & a_{1,i_2} & \cdots & a_{1,i_{k-1}} \\ a_{2,i_1} & a_{2,i_2} & \cdots & a_{2,i_{k-1}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1,i_1} & a_{k-1,i_2} & \cdots & a_{k-1,i_{k-1}} \end{vmatrix} = \gamma \neq 0.$$

Finalmente, $\forall\, j \in \{1, \ldots, n\}$, se tiene por hipótesis

$$0 = \begin{vmatrix} a_{1,i_1} & \cdots & a_{1,i_{k-1}} & a_{1,j} \\ a_{2,i_1} & \cdots & a_{2,i_{k-1}} & a_{2,j} \\ \vdots & & \vdots & \vdots \\ a_{k,i_1} & \cdots & a_{k,i_{k-1}} & a_{k,j} \end{vmatrix}$$

$$= \alpha_1 a_{1j} + \alpha_2 a_{2j} + \cdots + \gamma a_{k,j},$$

donde las α_s denotan los correspondientes menores a a_{sj} multiplicados por $(-1)^{(s+j)}$. Notése también que si $j = i_s$ para alguna $s \in \{i_1, i_2, \ldots, i_{k-1}\}$, se trata de una matriz con 2 columnas iguales. Por lo tanto

$$\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \gamma A_k = \overline{0},$$

y como $\gamma \neq 0$, se tiene que $A_1, A_2, \ldots, A_k$ son linealmente dependientes. □

Corolario 4.8.2 *Sea A una matriz, entonces el rango de A es $k \Leftrightarrow$ todas las submatrices de $r \times r$, $r > k$ tienen determinante 0, y A tiene una submatriz de $k \times k$ con determinante no nulo.*

Demostración.
$\Rightarrow$) Si el rango de A es k, entonces A tiene k renglones linealmente independientes. En general, cualquier conjunto de generadores de un subespacio vectorial contiene una base (ejercicio). Se sigue entonces del teorema anterior que existe una submatriz de $k \times k$ con determinante distinto de 0. También se sigue del mismo teorema que si $r > k$, cualquier submatriz de $r \times r$ tiene determinante 0, ya que cualesquiera r renglones son linealmente dependientes.
$\Leftarrow$) Se sigue también del teorema anterior que como existe una submatriz de $k \times k$ con determinante no nulo, los renglones involucrados son linealmente independientes, y que si todo submatriz de $r \times r$, donde $r > k$, tiene determinante 0, entonces cualesquiera r vectores renglón son linealmente dependientes. En otras palabras, el rango es k.

□

Corolario 4.8.3 *El rango de una matriz es igual a la dimensión del subespacio vectorial generado por las columnas.*

Demostración. Provisionalmente denotamos por $rango_1$, a la definición original y $rango_2$ a la dimensión del subespacio vectorial generado por las columnas. Se sigue entonces del Corolario 4.8.2 que

$$rango_2(A) = rango_1(A^t) = rango_1(A),$$

ya que este corolario permite definir rango sin mencionar columnas o renglones, sino solamente submatrices, y éstas y sus transpuestas tienen los mismos determinantes.

□

Corolario 4.8.4 *Una matriz de $n \times n$ es de rango $n \Leftrightarrow$ su determinante es distinto de* 0.

Demostración. Se sigue del Teorema 4.8.1 que los vectores renglón son linealmente dependientes si y sólo si el determinante de la matriz es 0.

□

EJERCICIO 4.8

1. Pruebe que cualquier conjunto finito de generadores de un subespacio vectorial en $\mathbb{R}^n$ contiene una base.

4.9. El determinante como área o volumen

Definición 58 *Sea* $f : \underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{m \text{ veces}} \longrightarrow \mathbb{R}$, *se dice que* f *es multilineal, si es lineal con respecto a cada variable, i.e., se cumplen las siguientes dos condiciones* $\forall\, \lambda \in \mathbb{R},\ \forall\, k \in \{1, 2, \ldots, m\}$ *y* $\forall\ x_k, x_{k'} \in \mathbb{R}^n$:

i) $f(x_1, \ldots, x_{k-1}, x_k + x_{k'}, x_{k+1}, \ldots, x_m)$

$= f(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_m) + f(x_1, \ldots, x_{k-1}, x_{k'}, x_{k+1}, \ldots, x_m),$

ii)

$$f(x_1, \ldots, \lambda\, x_k, \ldots, x_n) = \lambda\, f(x_1, \ldots, x_k, \ldots, x_n).$$

Obsérvese que la función determinante se puede pensar como una función multilineal de n vectores en $\mathbb{R}^n$, ya sea las columnas o los renglones (por las propiedades demostradas). Esto se ve claro al escribir, por ejemplo

$$Det\ A \ = \ Det\ \ (v_1, v_2, \ldots, v_n), \qquad \text{donde} \qquad A = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Recordamos que el determinante como función de los n vectores renglón (o columna) en $\mathbb{R}^n$ cumple las siguientes tres propiedades:

i) $(\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n) \xrightarrow{Det} \mathbb{R}$ es multilineal,

ii) $Det\ \ (v_1, v_2, \ldots, v_{j-1}, v_j, v_j, v_{j+2}, \ldots, v_n) = 0$, *i.e.*, si se tienen dos renglones consecutivos iguales, el determinante es 0,

iii) $Det\ \ (e_1, e_2, \ldots, e_n) = 1$, *i.e.*, $Det\ \ (I) = 1$.

Teorema 4.9.1 *Existe una única función con dominio*

$$\underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ veces}}$$

y valores reales, que satisface i), ii) y iii), esta función es el determinante.

Probamos primero un lema.

Lema 4.9.2 *Bajo las hipótesis del Teorema 4.9.1, una transposición de elementos consecutivos cambia el signo de la función.*

Demostración.

$$0 = f(v_1, \ldots, v_k + v_{k'}, v_k + v_{k'}, v_{k+2}, \ldots, v_n),$$

por lo cual

$$f(v_1, \ldots, v_{k-1}, v_k, v_{k'}, v_{k+2}, \ldots, v_n) = -f(v_1, \ldots, v_{k-1}, v_{k'}, v_k, v_{k+2}, \ldots, v_n).$$

□

Demostración del Teorema 4.9.1 Sea f dicha función, entonces

$$f\left(\sum_{i=1}^{n} a_{1\,i}\, e_i, \ldots, \sum_{i=1}^{n} a_{n\,i}\, e_i\right) = \sum_{\sigma \in S_n} a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)}\, f(e_{\sigma(1)}, \ldots e_{\sigma(n)}).$$

Ya que al aplicar la propiedad de multilinealidad se tienen tantos sumandos como permutaciones.

Finalmente, usando el Teorema 4.3.3, si la permutación

$$\sigma = \begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix}$$

consiste de m inversiones, entonces se puede mediante m transposiciones de elementos consecutivos, componer σ y obtener la identidad.

Por lo cual

$$f\left(\sum_{i=1}^{n} a_{1\,i}\, e_i, \ldots, \sum_{i=1}^{n} a_{n\,i}\, e_i\right) = \sum_{\sigma \in S_n} a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)}\, \varepsilon(\sigma).$$

□

Área - volumen

Definición 59 *Sean $v, w \in \mathbb{R}^2$ y $P(v, w)$ el paralelogramo generado por v y w. Se denota por $vol_0\,(v, w)$ al área de $P(v, w)$ si $Det\,(v, w) \geq 0$ y menos el área de $P(v, w)$ si $Det\,(v, w) \leq 0$.*

Nótese que si el área se denota por $vol\,(v,w)$, entonces

$$|vol_0\,(v,w)| \;=\; vol\,(v,w).$$

Se usarán propiedades simples del área como las siguientes:

1. *Área (segmento)*$= 0$,
2. Si $\acute{A}rea(A) = m$, $\acute{A}rea(T(A)) = m$, donde T es una traslación,
3. Si $A \cap B = \emptyset$ o $A \cap B$ es un conjunto de área cero, entonces

$$\acute{A}rea(A \cup B) = \acute{A}rea(A) \cup \acute{A}rea(B),$$

La prueba formal de estas propiedades se lleva a cabo en el curso de cálculo de varias variables, donde se define al área como una integral doble

$$A_E(R) = \int_R d\,A.$$

La invariabilidad bajo traslación se sigue del teorema de cambio de variable aplicado a $x \mapsto x + x_0$, ya que el jacobiano es la identidad.

Teorema 4.9.3 *Sean* $v, w \in \mathbb{R}^2$, *entonces*

$$vol_0\,(v,w) \;=\; Det\,(v,w).$$

Necesitamos primero unos lemas antes de probar el teorema.

Lema 4.9.4 *Si* $c \in \mathbb{R}$, $c \geq 0$ *y* $v, w \in \mathbb{R}^2$, *entonces*

$$vol\,(c\,v,w) \;=\; c\,vol\,(v,w).$$

Demostración. Basta probar el caso cuando v y w son linealmente independientes. En este caso, si $n \in \mathbb{N}$, el paralelogramo generado por $n\,v$ y w consiste de n paralelogramos paralelos a $P(v,w)$.

Los paralelogramos obtenidos son traslaciones por $v, 2\,v, \ldots, (n-1)\,v$ del paralelogramo $P\,(v,w)$ y se tiene que

$$vol\,(n\,v,w) \;=\; n\,vol\,(v,w).$$

También

$$vol\left(\frac{v}{n}, w\right) = \frac{1}{n} vol\,(v, w),$$

y si $m, n \in \mathbb{N}$

$$vol\left(\frac{m}{n} v, w\right) = \frac{m}{n} vol\,(v, w).$$

Esto se sigue, ya que si $v_1 = \frac{v}{n}$, se tiene de la observación anterior que

$$vol\,(n\, v_1, w) = n\; vol\,(v_1, w).$$

Más aún,

$$\begin{aligned} vol\left(m\,\frac{1}{n}\, v, w\right) &= m\; vol\,\left(\frac{1}{n}\, v, w\right) \\ &= m\,\frac{1}{n}\; vol\,(v, w). \end{aligned}$$

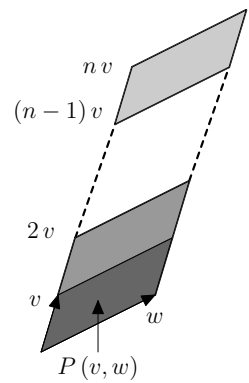


Figura 4.1: Paralelogramo generado por $n\, v$ y w

Finalmente, si $c \in \mathbb{R},\ c > 0$, entonces $vol\,(c\, v, w) = c\, vol\,(v, w)$. Para probar esto, sean $r_1, r_2 \in \mathbb{Q},\ 0 < r_1 < c < r_2$. Entonces

$$r_1\, vol\,(v, w) = vol\,(r_1\, v, w) \leq vol\,(c\, v, w) \leq vol\,(r_2\, v, w) = r_2\, vol\,(v, w),$$

(veáse la Figura 4.2). Haciendo tender r_1 y r_2 a c, en el límite se tiene que

$$c\, vol\,(v,w) \;=\; vol\,(c\,v,w).$$

□

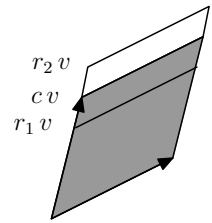


Figura 4.2: $c\, vol\,(v,w) \;=\; vol\,(c\,v,w)$

Por otro lado, al trasladar el paralelogramo $P\,(v,w)$ por $-v$ se obtiene el siguiente resultado.

Lema 4.9.5 *Sean* $v, w \in \mathbb{R}^2$, *entonces*

$$vol\,(-v,w) \;=\; vol\,(v,w).$$

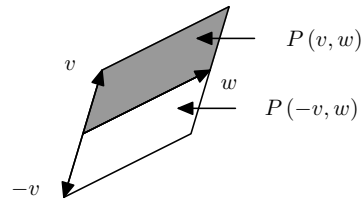


Figura 4.3: $vol\,(-v,w) \;=\; vol\,(v,w)$

Lema 4.9.6 *Sean* $v, w \in \mathbb{R}^2$, *y* $c \in \mathbb{R}$, *entonces*

$$vol_0\,(c\,v,w) \;=\; c\,vol_0\,(v,w).$$

Demostración. El resultado es evidente si v y w son linealmente dependientes. Consideramos casos

Caso 1: $Det\,(v,w) > 0$ *y* $c < 0$. En este caso, $Det\,(c\,v,w) \leq 0$ y tomando $c = -d$, se sigue en virtud del Lema 4.9.5 que

$$vol_0\,(c\,v,w) = -vol(c\,v,w) = -vol\,(-d\,v,w) = -vol\,(d\,v,w)$$

$$= -d\,vol\,(v,w) = c\,vol\,(v,w) = c\,vol_0\,(v,w).$$

El caso $Det\,(v,w) > 0,\ c > 0$ ya se probó. De manera análoga, se prueban los dos casos restantes (ejercicio). □

Para probar que vol_0 distribuye la suma usamos los siguientes resultados.

Lema 4.9.7 *Sean* $v, w \in \mathbb{R}^2$, *entonces*

$$vol_0\,(v,w) = -vol_0\,(w,v).$$

Demostración.

Caso 1: $Det\,(v,w) \leq 0$.

$$vol_0\,(v,w) = -vol\,(v,w) = -vol\,(w,v) = -vol_0\,(w,v),$$

puesto que $Det\,(w,v) \geq 0$.

Caso 2: $Det\,(v,w) > 0$.

$$vol_0\,(v,w) = vol\,(v,w) = vol\,(w,v) = -vol_0\,(w,v),$$

puesto que $Det\,(w,v) \leq 0$.

□

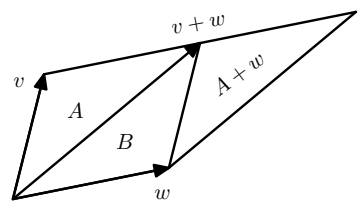


Figura 4.4: $vol\,(v+w,w) = vol\,(v,w)$

Lema 4.9.8 *Sean* v, w *vectores en el plano, entonces*

$$vol\,(v+w,w) = vol\,(v,w).$$

Demostración. El paralelogramo generado por v y w consiste de 2 triángulos A y B. Por otra parte, el generado por $v+w$ y w consiste del triángulo B y del triángulo $A+w$, por lo que tienen la misma área (ver la Figura 4.4).

□

Lema 4.9.9 *Sean* $v, w \in \mathbb{R}^2$, *y* $d \in \mathbb{R}$, *entonces*

$$vol_0\,(v + d\,w, w) \;=\; vol_0\,(v, w).$$

Demostración.

$$d\,vol_0\,(v + dw, w) \;=\; vol_0\,(v + dw, dw) \;=\; vol_0\,(v, dw) \;=\; d\,vol_0\,(v, w).$$

□

Demostración del Teorema 4.9.3 Basta probar que el $vol_0\,(v, w)$ cumple las propiedades *i*), *ii*) y *iii*). Hay que demostrar *i*), ya que *ii*) y *iii*) son triviales. Es decir, hay que probar la multilinealidad. En virtud de los Lemas 4.9.4, 4.9.6 y 4.9.7 es suficiente probar que

$$vol_0\,(v_1 + v_2, w) \;=\; vol_0\,(v_1, w) + vol_0\,(v_2, w).$$

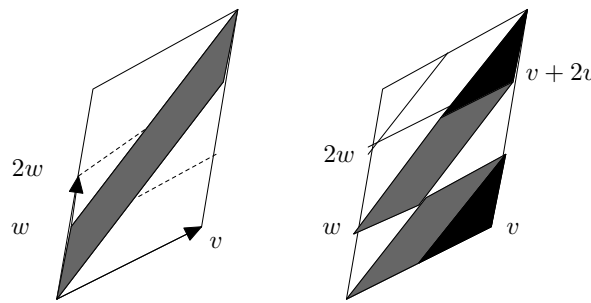


Figura 4.5: Interpretación geométrica del Lema 4.9.9, para el caso $d = 2$, se tiene que $vol\,(v + 2w, w) \;=\; vol\,(v, w)$

Si $w = 0$ no hay nada que probar. De otra manera, tomando como base $\{v, w\}$, se tiene

$$v_1 = c_1\,v + d_1\,w \quad \text{y} \quad v_2 = c_2\,v + d_2\,w.$$

Finalmente, usando los Lemas 4.9.6 y 4.9.9, se tiene

$$vol_0\,(v_1 + v_2, w) = vol_0\,(c_1\,v + d_1\,w + c_2\,v + d_2\,w, w)$$

$$= (c_1 + c_2)\;vol_0\,(v, w) = c_1\;vol_0\;(v, w) + c_2\;vol_0\;(v, w)$$

$$= vol_0\;(c_1\,v, w) + vol_0\;(c_2\,v, w)$$

$$= \;vol_0\;(v_1, w) + vol_0\;(v_2, w).$$

□

Un hecho interesante es que la misma prueba se aplica para $\mathbb{R}^3$ o $\mathbb{R}^n$. Esto se sigue ya que las condiciones que caracterizan el determinante involucran dos coordenadas a la vez, por lo cual los argumentos se pueden adaptar al caso bidimensional. Por ejemplo, tomando paralelepípedos y trasladándolos como en el caso bidimensional, se siguen todos los argumentos.

Nótese que el teorema se puede formular como sigue

$$|\,Det\,(v, w)\,| \;=\; vol\,(v, w).$$

Otros resultados interesantes sobre áreas y determinantes se puede consultar en [4]. Uno de estos resultados establece que si $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ es lineal, entonces

$$\acute{A}rea\;L(P) = |det\;A|\;\acute{A}rea\;P,$$

donde A es la matriz que define L.

EJERCICIO 4.9

1. Terminar la prueba del Lema 4.9.6.

Capítulo 5

Sistemas de ecuaciones lineales

5.1. Preliminares

Estudiaremos un sistema de m ecuaciones lineales con n incógnitas:

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ & \vdots & & & & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array}\right. \tag{5.1}$$

Las incógnitas o variables por resolver son $x_1, x_2, \ldots, x_n$, las constantes a_{ij}, b_j se toman generalmente en un campo como los reales o los complejos, o también en un anillo como los enteros. A los términos $b_1, b_2, \ldots, b_m$ se les llama libres.

A un sistema se le asocian dos matrices

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix},$$

a la primera se les llama matriz del sistema y a la segunda se le conoce como la matriz aumentada.

Obsérvese que si denotamos las columnas de la matriz por $C_1, C_2, \ldots, C_n$ y a B la columna de los términos libres, el sistema (5.1) se puede reescribir como

$$x_1C_1 + x_2C_2 + \cdots + x_nC_n = B. \tag{5.2}$$

Más aún, si $X = (x_1, x_2, \ldots, x_n)$ es solución, entonces $B \in < C_1, \ldots, C_n >$.

Por ejemplo, el sistema

$$\begin{array}{rcrcrcl} 2x & + & 3y & - & z & = & 5 \\ x & - & y & + & z & = & 2 \end{array}$$

se puede escribir como $x(2,1) + y(3,-1) + z(-1,1) = (5,2)$.

Definición 60 *Al sistema* (5.1) *se le llama homogéneo si*

$$B = (b_1, \ldots, b_m) = \overline{0}.$$

Algunos sistemas tienen muchas soluciones, otros exactamente una y algunos otros ninguna.

Por ejemplo,

$$\left\{ \begin{array}{rcrcl} x & + & y & = & 3 \\ x & - & y & = & 1 \end{array} \right.$$

tiene una única solución, ya que las ecuaciones implican $2x = 4$, y $y = 1$.

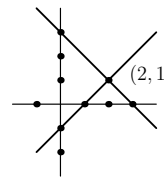


Figura 5.1: Las rectas $x + y = 3$, $x - y = 1$ se intersectan en $(2,1)$

Por otra parte, la ecuación $2x - y = 0$ tiene una infinidad de soluciones, éstas son los puntos de la recta $2x = y$.

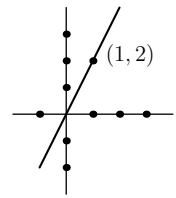


Figura 5.2: Las soluciones de $2x - y = 0$

Sin embargo, el sistema

$$\begin{cases} -x & + & y & = & -1 \\ -x & + & y & = & -2 \end{cases}$$

no tiene soluciones, de lo contrario, restando se tendría $0 = -1$.

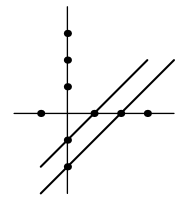


Figura 5.3: Las rectas $y = x - 1,\ y = x - 2$ no se intersecan, al ser paralelas

Ciertamente un sistema homogéneo siempre tiene al menos la solución trivial, *i.e.*, el vector $\overline{0}$.

5.2. Existencia de soluciones

Conocer el rango de las matrices del sistema es muy útil para resolver un sistema. En el siguiente teorema usamos la notación de la ecuación (5.2).

Teorema 5.2.1 *Un sistema de ecuaciones lineales tiene solución si y sólo si el rango de la matriz del sistema es igual al rango de la matriz aumentada.*

Demostración. La existencia de soluciones significa que existe $\overline{x} \in \mathbb{R}^n$, $\overline{x} = (x_1, \ldots, x_n)$ tal que

$$x_1\, C_1 + x_2\, C_2 + \cdots + x_n\, C_n = B.$$

Esto sucede si y sólo si

$$B \in < C_1, C_2, \ldots, C_n > .$$

Más aún, esta condición se cumple si y sólo si

$$< C_1, C_2, \ldots, C_n > = < B, C_1, C_2, \ldots, C_n > .$$

Finalmente, esta condición equivale a que los rangos de la matriz del sistema, y de la matriz aumentada sean iguales: si $B \notin < C_1, C_2, \ldots, C_n >$, entonces el rango de $< B, C_1, C_2, \ldots, C_n >$ es mayor al de la matriz del sistema, y B junto con una base de $\{C_1, C_2, \ldots, C_n\}$ forman un conjunto linealmente independiente.

□

Ejemplos

1. Consideremos el siguiente sistema de ecuaciones

$$\begin{cases} x + y - w = 0 \\ x + z - w = -1 \\ -x + y - 2z + w = 3\,. \end{cases}$$

La matriz aumentada es $\begin{pmatrix} 1 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & -1 & -1 \\ -1 & 1 & -2 & 1 & 3 \end{pmatrix}$

$$\sim \begin{pmatrix} 1 & 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 2 & -2 & 0 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

y la matriz tiene rango 3. Las mismas operaciones muestran que la matriz del sistema

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \\ -1 & 1 & -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

tiene rango 2, por lo que el sistema no tiene soluciones.

2. Analizamos ahora el sistema

$$\begin{cases} x - y + 2z = 1 \\ y - z = 1 \\ 3x + y - z = 0 \\ 4x + y = 2\,. \end{cases}$$

La matriz aumentada de este sistema de ecuaciones es

$$\begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 3 & 1 & -1 & 0 \\ 4 & 1 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 4 & -7 & -3 \\ 0 & 5 & -8 & -2 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -3 & -7 \\ 0 & 0 & -3 & -7 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -3 & -7 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

es decir, tiene rango 3. Las mismas operaciones establecen que

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \\ 3 & 1 & -1 \\ 4 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix},$$

que también tiene rango 3, y por lo tanto el sistema sí tiene solución.

3. Consideramos el sistema $x_1\, C_1 + x_2\, C_2 + \cdots + x_n\, C_n = B,$ donde los vectores columna C_j son vectores en $\mathbb{R}^n$ (y B también), supóngase también que

$$det(C_1 C_2 \cdots C_n) \neq 0.$$

Entonces el sistema tiene solución, esto se sigue ya que esta hipótesis implica que la matriz del sistema tiene rango n (Corolario 4.8.4). Además, el rango de la matriz aumentada (que siempre es mayor o igual que el de la matriz del sistema) es $\leq n,$ ya que sus vectores columna viven en $\mathbb{R}^n$. Es decir, ambos rangos son iguales.

4. Otra prueba de que cualquier sistema homogéneo tiene solución es observando que las columnas de las dos matrices del sistema generan el mismo subespacio vectorial (la última columna es de ceros).

5. Un sistema de m ecuaciones y n incógnitas, donde $m < n,$ y tal que el rango de la matriz sea m, tiene solución. Esto se sigue de que el rango de la matriz aumentada es $\leq m$ (los vectores columna viven en $\mathbb{R}^m$), y por lo tanto es m.

EJERCICIO 5.2

1. Determine si los siguientes sistemas tienen solución.

$$a)\begin{cases} x - y + 2z = 7 \\ y + z = 5 \\ 2z = 6, \end{cases} \qquad b)\begin{cases} x + y - z = 1 \\ 2x - 3z = 5 \\ 2y + 5z = 2, \end{cases}$$

$$c)\begin{cases} 2x + 3y + z = 7 \\ 2x - 4y + 6z = 10 \\ 3x - 5y + 3z = 4, \end{cases} \qquad d)\begin{cases} x - y + 2z = -2 \\ 3x - 3y + 6z = 1 \\ 2x - 2y + 4z = 0. \end{cases}$$

5.3. Sistemas de n ecuaciones y n incógnitas

Consideremos el siguiente sistema de ecuaciones

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = k_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = k_2 \\ \vdots \qquad\qquad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = k_n. \end{cases} \tag{5.3}$$

Denotaremos por d al determinante de la matriz del sistema, *i.e.*,

$$d = det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Obsérvese que si $d \neq 0$, el sistema (5.3) tiene solución. Esto se sigue del argumento del ejemplo 3 de la sección anterior. Para encontrar la solución (se probará que es única) se usan los siguientes determinantes

$$d_1 = \begin{vmatrix} k_1 & a_{12} & \cdots & a_{1n} \\ k_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots, \\ k_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \quad d_2 = \begin{vmatrix} a_{11} & k_1 & a_{13} & \cdots & a_{1n} \\ a_{21} & k_2 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & k_n & a_{n3} & \cdots & a_{nn} \end{vmatrix},$$

$$\cdots, \qquad d_n = \begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & k_1 \\ a_{21} & \cdots & a_{2,n-1} & k_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & k_n \end{vmatrix}.$$

Ahora sea $S = (s_1, s_2, \ldots, s_n)$ una solución del sistema, entonces

$$\begin{array}{ccccccc} a_{11}s_1 & + & \cdots & + & a_{1n}s_n & = & k_1 \\ a_{21}s_1 & + & \cdots & + & a_{2n}s_n & = & k_2 \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{n1}s_1 & + & \cdots & + & a_{nn}s_n & = & k_n, \end{array}$$

y usando las propiedades del determinante se tiene que

$$d_1 = \begin{vmatrix} k_1 & a_{12} & \cdots & a_{1n} \\ k_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11}s_1 + \cdots + a_{1n}s_n & a_{12} & \cdots & a_{1n} \\ a_{21}s_1 + \cdots + a_{2n}s_n & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}s_1 + \cdots + a_{nn}s_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11}s_1 & a_{12} & \cdots & a_{1n} \\ a_{21}s_1 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}s_1 & a_{n2} & \cdots & a_{nn} \end{vmatrix} = s_1 d.$$

Análogamente, $d_2 = s_2 d, \ldots, d_n = s_n d$. Hemos probado el siguiente teorema.

Teorema 5.3.1 *(Regla de Cramer) El sistema* (5.3) *tiene una única solución*

$$s = \left(\frac{d_1}{d}, \frac{d_2}{d}, \ldots, \frac{d_n}{d} \right).$$

Corolario 5.3.2 *Si la matriz de un sistema homogéneo de n ecuaciones y n incógnitas, tiene determinante distinto de 0, entonces su única solución es la trivial, i.e., el vector* $\overline{0}$.

Demostración. Se tiene que $d_1 = \cdots = d_n = 0$, ya que estas matrices tienen una columna de ceros. □

Otra demostración se obtiene al observar que $\overline{0}$ siempre es solución, y por la regla de Cramer es la única.

Ejemplos

1. Dado el sistema

$$\left\{\begin{array}{rcrcl} x & + & y & = & a \\ x & - & y & = & b \end{array}\right., \qquad \text{se tiene que} \quad d = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2,$$

$$d_1 = \begin{vmatrix} a & 1 \\ b & -1 \end{vmatrix} = -a - b$$

$$\text{y} \quad d_2 = \begin{vmatrix} 1 & a \\ 1 & b \end{vmatrix} = b - a,$$

$\therefore$ la solución única es $(x, y) = \left(\frac{a+b}{2}, \frac{a-b}{2}\right)$.

2. Para el sistema

$$\left\{\begin{array}{rcrcl} x \operatorname{sen}\alpha & + & y\cos\alpha & = & \operatorname{sen} 2\alpha \\ x\cos\alpha & - & y \operatorname{sen}\alpha & = & \cos 2\alpha, \end{array}\right.$$

$$\text{se tiene que} \quad d = \begin{vmatrix} \operatorname{sen}\alpha & \cos\alpha \\ \cos\alpha & -\operatorname{sen}\alpha \end{vmatrix} = -1,$$

$$\begin{aligned} d_1 &= \begin{vmatrix} \operatorname{sen} 2\alpha & \cos\alpha \\ \cos 2\alpha & -\operatorname{sen}\alpha \end{vmatrix} \\ &= -\operatorname{sen}\alpha \operatorname{sen} 2\alpha - \cos\alpha\cos 2\alpha \\ &= -\operatorname{sen}\alpha(2\operatorname{sen}\alpha\cos\alpha) - \cos\alpha(\cos^2\alpha - \operatorname{sen}^2\alpha) \\ &= -2\operatorname{sen}^2\alpha\cos\alpha - \cos\alpha(1 - 2\operatorname{sen}^2\alpha) \\ &= -\cos\alpha, \end{aligned}$$

$$\begin{aligned} d_2 &= \begin{vmatrix} \operatorname{sen}\alpha & \operatorname{sen} 2\alpha \\ \cos\alpha & \cos 2\alpha \end{vmatrix} \\ &= (\cos^2\alpha - \operatorname{sen}^2\alpha)\operatorname{sen}\alpha - 2\operatorname{sen}\alpha\cos^2\alpha \\ &= (2\cos^2\alpha - 1)\operatorname{sen}\alpha - 2\operatorname{sen}\alpha\cos^2\alpha \\ &= -\operatorname{sen}\alpha, \end{aligned}$$

y la solución es $(\cos\alpha, \operatorname{sen}\alpha)$.

Usando un poco de geometría analítica y notación compleja el último ejemplo se puede interpretar geométricamente.

Se busca $(x, y) \in \mathbb{R}^2$ tal que

$$\begin{pmatrix} \operatorname{sen}\alpha & \cos\alpha \\ \cos\alpha & -\operatorname{sen}\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \operatorname{sen} 2\alpha \\ \cos\ 2\alpha \end{pmatrix},$$

obsérvese que

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \operatorname{sen}\alpha & \cos\alpha \\ \cos\alpha & -\operatorname{sen}\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha. \end{pmatrix}. \tag{5.4}$$

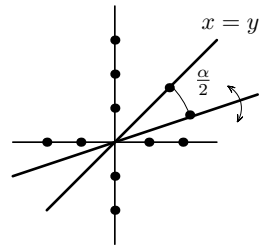


Figura 5.4: Reflexión en la recta por 0 y $e^{i(\frac{\pi}{4}-\frac{\alpha}{2})}$

Ahora las transformaciones definidas por

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \text{ y } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

son, respectivamente, rotación por α, y reflexión en la recta $x = y$ (la última intercambia x por y). Si denotamos la rotación por R y la reflexión por T, usando (5.4) se tiene

$$\begin{pmatrix} \operatorname{sen}\alpha & \cos\alpha \\ \cos\alpha & -\operatorname{sen}\alpha \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}, \tag{5.5}$$

(ya que $T = T^{-1}$) y como $R(e^{i\alpha}) = e^{i(2\alpha)} = \cos 2\alpha + i\operatorname{sen} 2\ \alpha$, al componer con T se tiene $(\operatorname{sen} 2\alpha, \cos 2\alpha)$, *i.e.*, la solución es $(\cos\alpha, \operatorname{sen}\alpha)$.

Obsérvese que se sigue de la ecuación (5.5) que la transformación definida por

$$\begin{pmatrix} \operatorname{sen}\alpha & \cos\alpha \\ \cos\alpha & -\operatorname{sen}\alpha \end{pmatrix}$$

es la reflexión en la recta que pasa por $e^{i(\frac{\pi}{4}-\frac{\alpha}{2})}$, ya que al rotar $e^{i(\frac{\pi}{4}-\frac{\alpha}{2})}$ por $e^{i\alpha}$ y reflejar en $x=y$, se tiene $e^{i(\frac{\pi}{4}-\frac{\alpha}{2})}$ (se sigue del álgebra lineal que como la matriz tiene determinante -1 es una reflexión). Véase la Figura 5.4.

El símbolo e^{ix} corresponde a $(\cos x, \operatorname{sen} x)$. La multiplicación de matrices se define como

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} a\,\alpha + b\,\gamma & a\,\beta + b\,\delta \\ c\,\alpha + d\,\gamma & c\,\beta + d\,\delta \end{pmatrix},$$

y

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}.$$

Otros ejemplos

1. Se resuelve el sistema

$$\begin{cases} x + y + z = 1 \\ ax + by + cz = k \\ a^2x + b^2y + c^2z = k^2, \end{cases}$$

donde a, b, c son tres números distintos. El determinante del sistema es de tipo Van der Monde y está dado por

$$d = \begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (c-b)(b-a)(c-a).$$

Los determinantes d_1, d_2, d_3 son también de tipo Van der Monde, y se sigue de la regla de Cramer que la solución está dada por

$$\left(\frac{(c-k)(c-b)(b-k)}{(c-b)(b-a)(c-a)}, \frac{(c-k)(k-a)(c-a)}{(c-b)(b-a)(c-a)}, \frac{(k-b)(k-a)(b-a)}{(c-b)(b-a)(c-a)}\right)$$

$$= \left(\frac{(c-k)(b-k)}{(b-a)(c-a)}, \frac{(c-k)(k-a)}{(c-b)(b-a)}, \frac{(k-b)(k-a)}{(c-b)(c-a)}\right).$$

2. Exprese $K=(5,1,11)$ como combinación lineal de $B_1=(3,2,2)$, $B_2=(2,3,1)$ y $B_3=(1,1,3)$. Se debe resolver

$$x_1B_1 + x_2B_2 + x_3B_3 = K,$$

es decir,

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 5 \\ 2x_1 + 3x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 11. \end{cases}$$

Se tiene

$$d = \begin{vmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 6 & 2 & 1 \\ 6 & 3 & 1 \\ 6 & 1 & 3 \end{vmatrix} = 6 \begin{vmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{vmatrix} = 6 \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 2 \end{vmatrix} = 12,$$

$\therefore$ hay una única solución.

Ahora

$$d_1 = \begin{vmatrix} 5 & 2 & 1 \\ 1 & 3 & 1 \\ 11 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 2 \\ 1 & 1 & 3 \\ 3 & 11 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 2 \\ 0 & -4 & 1 \\ 0 & -4 & -5 \end{vmatrix} = 24,$$

$$d_2 = \begin{vmatrix} 3 & 5 & 1 \\ 2 & 1 & 1 \\ 2 & 11 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 \\ 1 & 2 & 1 \\ 3 & 2 & 11 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 \\ 0 & -1 & -4 \\ 0 & -7 & -4 \end{vmatrix} = -24,$$

$$d_3 = \begin{vmatrix} 3 & 2 & 5 \\ 2 & 3 & 1 \\ 2 & 1 & 11 \end{vmatrix} = 3 \begin{vmatrix} 3 & 1 \\ 1 & 11 \end{vmatrix} - 2 \begin{vmatrix} 2 & 1 \\ 2 & 11 \end{vmatrix} + 5 \begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix}$$
$$= 3 \cdot 32 - 40 - 20 = 36,$$

$\therefore$ la solución es $(2, -2, 3)$.

3. ¿Son los vectores $B_1 = (1, 5, 3)$, $B_2 = (2, 1, -1)$ y $B_3 = (4, 2, 1)$ linealmente independientes?

Se debe probar que el determinante de la matriz formada con estos tres vectores es distinto de 0, esto se sigue del teorema que relaciona el rango con vectores linealmente independientes. Dicho de otra manera, escribiendo $x_1 B_1 + x_2 B_2 + x_3 B_3 = \overline{0}$, la independencia lineal se sigue si x_1, x_2, x_3 son 0, esto es, la única solución es la trivial, lo cual usando el teorema de Cramer se obtiene si el determinante es distinto de 0. En efecto, al calcular se obtiene

$$\begin{vmatrix} 1 & 2 & 4 \\ 5 & 1 & 2 \\ 3 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 0 & -9 & -18 \\ 0 & -7 & -11 \end{vmatrix} = -9 \begin{vmatrix} 1 & 2 \\ -7 & -11 \end{vmatrix} = -27 \neq 0.$$

EJERCICIOS 5.3

1. Encuentre la solución a los siguientes sistemas usando la regla de Cramer.

$$a)\begin{cases} x + y - z = -2 \\ 3x - y + 2z = 4 \\ -x + 2y - z = 6, \end{cases} \qquad b)\begin{cases} x - y + z = -1 \\ 2x - y - z = 4 \\ 3x + 2y + z = -1, \end{cases}$$

$$c)\begin{cases} x - y + 3z = 4 \\ 2x - y + z = 6 \\ 3x - 2y + 2z = 10. \end{cases}$$

2. Exprese el vector $K = (1, 5, 9)$ como combinación lineal de los vectores $B_1 = (1, 3, 7)$, $B_2 = (1, 2, 3)$ y $B_3 = (0, 1, 1)$.

5.4. Sistemas homogéneos, funciones lineales

Los sistemas homogéneos son aquéllos de la forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \qquad\vdots \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0, \end{cases} \tag{5.6}$$

o en la notación de la ecuación (5.2) $x_1C_1 + x_2C_2 + \cdots + x_nC_n = \overline{0}$.

Teorema 5.4.1 *Si W es el conjunto de soluciones de* (5.6)*, entonces W es un subespacio vectorial de $\mathbb{R}^n$.*

Demostración. Evidentemente $\overline{0} \in W$. También, si $X, Y \in W$, donde $X = (x_1, x_2, \ldots, x_n)$ y $Y = (y_1, y_2, \ldots, y_n)$, se tiene que

$$\begin{aligned} x_1\,C_1 + \cdots + x_n\,C_n &= \overline{0} \\ y_1\,C_1 + \cdots + y_n\,C_n &= \overline{0}, \end{aligned}$$

y en consecuencia $(x_1 + y_1)\,C_1 + \cdots + (x_n + y_n)\,C_n = \overline{0}$, *i.e.*, $X + Y \in W$.

Finalmente, si $\lambda \in \mathbb{R}$ y $X = (x_1, x_2, \ldots, x_n) \in W$, entonces

$$\lambda\,x_1C_1 + \cdots + \lambda\,x_nC_n = \overline{0}$$

y $\lambda X \in W$. □

Obsérvese que en general para cualquier sistema de ecuaciones, como en (5.1), siempre se pueden reordenar las variables, renombrándolas de tal manera que las primeras columnas $C_1, C_2, \ldots, C_r$ sean una base del espacio vectorial generado por las columnas, *i.e.*, $< C_1, C_2, \ldots, C_n >$. En el caso homogéneo probaremos que W, el espacio de soluciones del sistema, tiene dimensión $n - r$.

Los siguientes resultados relacionan la dimensión de W y n.

Teorema 5.4.2 *Sean $C_1, C_2, \ldots, C_r$ una base del subespacio vectorial W generado por las columnas $C_1, C_2, \ldots, C_n$ del sistema* (5.2)*, i.e.,*

$$x_1 C_1 + x_2 C_2 + \cdots + x_n C_n = B.$$

Supóngase también que el sistema tiene solución, entonces dados $n-r$ números

$$s_{r+1}, s_{r+2}, \ldots, s_n,$$

existen otros r números únicos $s_1, s_2, \ldots, s_r$, tales que

$$s = (s_1, s_2, \ldots, s_r, s_{r+1}, \ldots, s_n)$$

es solución de (5.2)*.*

Demostración. Sea $C = s_{r+1}C_{r+1} + s_{r+2}C_{r+2} + \cdots + s_n C_n$, como $C_1, C_2, \ldots, C_r$ generan W, entonces $C \in W$. Además, como el sistema tiene solución se tiene que $B \in W$. Por lo cual $B - C \in W$ y existen $s_1, s_2, \ldots, s_r \ni$

$$B - C = s_1 C_1 + \cdots + s_r C_r,$$

i.e.,

$$B = s_1 C_1 + \cdots + s_r C_r + s_{r+1} C_{r+1} + \cdots + s_n C_n,$$

y $s = (s_1, s_2, \ldots, s_r, s_{r+1}, \ldots, s_n)$ es solución.

Para probar la unicidad, supóngase que $(s_1', \ldots, s_r', s_{r+1}, \ldots, s_n)$ también es solución, entonces

$$s_1' \, C_1 + \cdots + s_r' \, C_r + s_{r+1} \, C_{r+1} + \cdots + s_n \, C_n \, = \, B$$

y

$$s_1 \, C_1 + \cdots + s_r \, C_r + s_{r+1} \, C_{r+1} + \cdots + s_n \, C_n, = \, B.$$

Por lo cual restando, se tiene que

$$(s_1^{'} - s_1)\, C_1 + \cdots + (s_r^{'} - s_r)\, C_r = \overline{0},$$

y como $C_1, C_2, \ldots, C_r$ son linealmente independientes, se sigue que $s_i^{'} = s_i$ $\forall\, i \in \{1, \ldots, r\}$.

□

Definición 61 *Se dice que una función* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *es lineal, si cumple las siguientes dos condiciones*

a) $f(X+Y) = f(X) + f(Y) \qquad \forall\, X, Y \in \mathbb{R}^n,$

b) $f(t\,X) = t\, f(X) \qquad \forall\, t \in \mathbb{R} \quad y \quad \forall\, X \in \mathbb{R}^n.$

Nótese que si f es lineal, entonces $f(\overline{0}) = \overline{0}$. Esto es consecuencia de la condición b), puesto que $f(0X) = 0f(X)$.

Si $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es lineal, al conjunto $\{X \in \mathbb{R}^n \mid f(X) = 0\}$ se le llama el núcleo de f y se le denota por $ker\, f$ (del alemán kernel). Obsérvese que se sigue de manera inmediata de la definición que $ker\, f$ es un subespacio vectorial de $\mathbb{R}^n$.

Por otra parte, a un sistema de ecuaciones como (5.2) con n incógnitas y m ecuaciones

$$x_1 C_1 + \cdots + x_n C_n = B,$$

se le puede asociar una función f de $\mathbb{R}^n$ en $\mathbb{R}^m$ con regla de correspondencia

$$f\left((x_1, \ldots, x_n)\right) \;=\; x_1 C_1 + \cdots + x_n C_n. \tag{5.7}$$

Esta función es lineal:

si $X = (x_1, \ldots, x_n)$ y $Y = (y_1, \ldots, y_n)$, entonces

$$\begin{aligned} f(X+Y) &= (x_1 + y_1)\, C_1 + \cdots + (x_n + y_n)\, C_n \\ &= x_1\, C_1 + \cdots + x_n\, C_n + y_1\, C_1 + \cdots + y_n\, C_n \\ &= f(X) + f(Y), \end{aligned}$$

y si $\lambda \in \mathbb{R}$ y $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$, entonces

$$f(\lambda\, X) = \lambda\, x_1\, C_1 + \cdots + \lambda\, x_n\, C_n = \lambda\, f(X).$$

Es importante destacar que $ker\, f$ es el conjunto de soluciones del sistema homogéneo asociado, es decir, el definido por la ecuación

$$x_1 C_1 + \cdots + x_n C_n = \overline{0}.$$

Lema 5.4.3 *Sea* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *lineal, entonces*

$$Im\, f = \{Y \in \mathbb{R}^m \mid \exists\, X \in \mathbb{R}^n \ \ni\ f(X) = Y\}$$

es un subespacio vectorial de $\mathbb{R}^m$.

Demostración. Si $Y_1, Y_2 \in Im\, f$, existen $X_1, X_2 \in \mathbb{R}^n \ \ni\ f(X_i) = Y_i$, $i = 1, 2$, por lo que

$$f(X_1 + X_2) = f(X_1) + f(X_2) = Y_1 + Y_2,$$

y $Y_1 + Y_2 \in Im\, f$,

También, si $Y \in Im\, f$ y $\lambda \in \mathbb{R}$, entonces existe $X \in \mathbb{R}^n \ \ni\ f(X) = Y$. Por lo que $f(\lambda X) = \lambda\, f(X) = \lambda\, Y$, y $\lambda Y \in Im\, f$. □

Nótese que en el caso de una función lineal f asociada a un sistema de la forma

$$x_1\, C_1 + \cdots + x_n\, C_n = B,$$

se tiene que $Im\, f$ es precisamente el subespacio vectorial generado por las columnas $< C_1, C_2, \ldots, C_n >$.

Teorema 5.4.4 *Sea* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *lineal, entonces*

$$dim(ker\, f) + dim(Im\, f) = n.$$

Demostración. Sean $C_1, \ldots, C_r$ una base de $Im\, f$ y $B_1, \ldots, B_k$ una base de $ker\, f$ (si $ker\, f = \{\overline{0}\}$ no se considera ningún vector B_i).

Ahora, existen vectores $D_1, \ldots, D_r \in \mathbb{R}^n \ \ni\ f(D_i) = C_i \ \ \forall\, i$. Se afirma que el conjunto de vectores $D_1, D_2, \ldots, D_r, B_1, B_2, \ldots, B_k$ constituyen una base de $\mathbb{R}^n$, lo cual prueba el teorema, por lo que basta probar la afirmación.

Los vectores son linealmente independientes ya que si

$$\lambda_1\, D_1 + \cdots + \lambda_r\, D_r + \mu_1\, B_1 + \cdots + \mu_k\, B_k = \overline{0},$$

entonces

$$f(\lambda_1\, D_1 + \cdots + \lambda_r\, D_r + \mu_1\, B_1 + \cdots + \mu_k\, B_k) = f(\overline{0}) = \overline{0}.$$

Por lo cual, usando la linealidad se tiene

$$\lambda_1\, f(D_1) + \cdots + \lambda_r\, f(D_r) = \overline{0} \quad \text{y} \quad \lambda_1\, C_1 + \cdots + \lambda_r\, C_r = \overline{0}.$$

Como $C_1, \ldots, C_r$ son linealmente independientes, $\lambda_1 = \cdots = \lambda_r = 0$, lo cual implica que

$$\mu_1 B_1 + \cdots + \mu_k B_k = \overline{0},$$

y entonces $\mu_1 = \cdots = \mu_k = 0$.

Probamos ahora que estos vectores también generan. Si $X \in \mathbb{R}^n$,

$$f(X) = \alpha_1 C_1 + \alpha_2 C_2 + \cdots + \alpha_r C_r,$$

en términos de los vectores D_i, $i = 1, 2, \ldots, r$, esto se escribe

$$f(X) = \alpha_1 f(D_1) + \alpha_2 f(D_2) + \cdots + \alpha_r f(D_r),$$

y la linealidad nos permite afirmar que

$$f(X - \alpha_1 D_1 - \cdots - \alpha_r D_r) = \overline{0}.$$

En consecuencia, $X - \alpha_1 D_1 - \cdots - \alpha_r D_r \in ker\, f$ y

$$X - \alpha_1 D_1 - \cdots - \alpha_r D_r = \mu_1 B_1 + \cdots + \mu_k B_k,$$

para algunas $\mu_j \in \mathbb{R}$, $j = 1, 2, \ldots, k$. Es decir, $D_1, \ldots D_r, B_1, \ldots, B_k$ generan $\mathbb{R}^n$. □

Este resultado implica que si se tiene una ecuación vectorial

$$x_1 C_1 + \cdots + x_n C_n = 0$$

y las columnas C_1,...,C_r son una base del subespacio vectorial $<C_1,\ldots,C_n>$, entonces

$$dim(ker\, f) = n - r,$$

donde f es como en (5.7). Esto es, el espacio de soluciones de un sistema homogéneo tiene dimensión $n - r$ (compárese este resultado con el Teorema 5.4.2).

5.5. Sistema homogéneo asociado

De manera natural al sistema (5.2)

$$x_1 C_1 + \cdots + x_n C_n = B,$$

se le asocia el sistema homogéneo (5.6)

$$x_1 C_1 + \cdots + x_n C_n = \overline{0},$$

y se le llama sistema homogéneo asociado.

Teorema 5.5.1 *Las soluciones del sistema* (5.2) *consisten en todos los vectores de la forma*

$$X_0 + S,$$

donde X_0 *es una solución particular de* (5.2) *y* S *es cualquier solución de* (5.6) *el sistema homogéneo asociado.*

Demostración. Sea $X_0 = (x_1, \ldots, x_n)$ y Y cualquier solución de (5.2), $Y = (y_1, \ldots, y_n)$, entonces

$$x_1\, C_1 + \cdots + x_n\, C_n = B$$

y

$$y_1\, C_1 + \cdots + y_n\, C_n = B,$$

por lo cual

$$(y_1 - x_1)\, C_1 + \cdots + (y_n - x_n)\, C_n = B - B = \overline{0},$$

y $Y - X_0 = S$ es solución de (5.6), *i.e.*, $Y = X_0 + S$.

Finalmente si S es solución de (5.6), $S = (s_1, \ldots, s_n)$, entonces

$$\begin{aligned}(x_1 + s_1)\, C_1 + \cdots + (x_n + s_n)\, C_n &= x_1\, C_1 + \cdots + x_n\, C_n + s_1\, C_1 + \cdots + s_n\, C_n \\ &= B + \overline{0} = B,\end{aligned}$$

y $X_0 + S$ es solución. □

Ejemplo

$$\text{Consideremos} \qquad x + y = 2.$$

El sistema homogéneo asociado es $x + y = 0$, una solución particular del sistema es $(1, 1)$, por lo que las soluciones de este sistema consisten de los puntos de la recta $y = -x$ trasladados por $(1, 1)$.

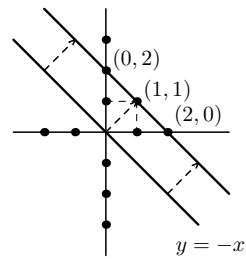


Figura 5.5: Soluciones a los sistemas
$x + y = 2$ y $x + y = 0$

5.6. Resolución de sistemas

La primera simplificación consiste en aplicar operaciones elementales a los vectores renglón.

Teorema 5.6.1 *Si un sistema de ecuaciones lineales se obtiene de otro mediante operaciones elementales en los renglones de la matriz aumentada, entonces ambos sistemas tienen las mismas soluciones.*

Antes de probar el teorema primero exhibimos un ejemplo, consideremos los sistemas:

$$(1)\ \begin{cases} x - 2y = 3 \\ 3x + 5y = 6 \end{cases} \qquad (2)\ \begin{cases} x - 2y = 3 \\ 9x + 15y = 18 \end{cases}$$

$$(3)\ \begin{cases} x - 2y = 3 \\ (3x+x) + (5y-2y) = 6+3 \end{cases} \qquad (4)\ \begin{cases} 3x + 5y = 6 \\ x - 2y = 3 \end{cases}$$

El sistema (2) se obtuvo de (1), multiplicando por 3 el 2o renglón. Además, multiplicando por 3 (o dividiendo entre 3), se sigue que

$$(x,y) \text{ es solución de } (1) \iff \text{es solución de } (2).$$

Ahora si (x,y) satisface (1) también satisface (3), ya que sumando las dos ecuaciones en (1) se obtiene la 2a ecuación en (3). Asimismo, si (x,y) satisface (3), restando en este sistema la 1a ecuación de la 2a, se sigue que (x,y) satisface (1). La equivalencia de (1) y (4) es evidente.

Demostración del Teorema 5.6.1 La 1a operación elemental de intercambiar renglones ciertamente no altera las soluciones del sistema, ya que si $(x_1, x_2, \ldots, x_n)$ cumple m ecuaciones, el orden en el que aparezcan es irrelevante.

Ahora si se multiplica en el sistema (5.1) el renglón $(a_{i1}, a_{i2}, \ldots, a_{in}, b_i)$ por $\lambda \neq 0$, entonces

$$\lambda\, x_1\, a_{i1} + \lambda\, x_2\, a_{i2} + \cdots + \lambda\, x_n\, a_{in} = \lambda\, b_i$$

$\iff$

$$x_1\, a_{i1} + x_2\, a_{i2} + \cdots + x_n\, a_{in} = b_i.$$

Finalmente, si $(x_1, \ldots, x_n)$ satisface

$$\begin{cases} x_1\, a_{i1} + x_2\, a_{i2} + \cdots + x_n\, a_{in} = b_i & (1) \\ x_1\, a_{j1} + x_2\, a_{j2} + \cdots + x_n\, a_{jn} = b_j, & (2) \end{cases}$$

entonces $(x_1, x_2, \ldots, x_n)$ satisface

$$x_1\,(a_{i1} + a_{j1}) + x_2\,(a_{i2} + a_{j2}) + \cdots + x_n\,(a_{in} + a_{jn}) = b_i + b_j. \qquad (3)$$

Y viceversa, si $(x_1, x_2, \ldots, x_n)$ satisface (3) y (2), entonces restando (2) de (3), se sigue que este punto satisface también la ecuación (1). □

Sabiendo que un sistema de ecuaciones lineales tiene solución, por ejemplo, verificando que los rangos de las matrices sean iguales, un método para resolverlo es el siguiente:

Paso 1. Aplicar operaciones elementales a los renglones de la matriz aumentada hasta que ésta sea escalonada.

Paso 2. Se pueden reordenar las columnas de la matriz del sistema (si es necesario), renombrando las variables, para que el nuevo sistema sea de la forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r + \cdots + a_{1n}x_n = b_1 \\ \qquad\quad a_{22}x_2 + \cdots + a_{2r}x_r + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\quad \vdots \qquad\qquad\quad \vdots \\ \qquad\qquad\qquad\quad a_{rr}x_r + \cdots + a_{rn}x_n = b_r, \end{cases} \qquad (5.8)$$

donde los primeros r vectores columna de la matriz del sistema son linealmente independientes.

Usando el Teorema 5.4.2 se obtienen todas las soluciones. Primero se aplica, para encontrar una particular. Posteriormente, se usa para encontrar las soluciones del sistema homogéneo asociado. Y de esta manera se obtienen todas las soluciones del sistema en virtud del Teorema 5.5.1. Recordamos que el Teorema 5.4.2 establece que dados $n - r$ números arbitrarios $s_{r+1}, s_{r+2}, \ldots, s_n$, existen otros r números únicos $s_1, s_2, \ldots, s_r$, tales que

$$s = (s_1, \ldots, s_r, s_{r+1}, \ldots, s_n)$$

es solución de (5.8).

Paso 3. Para encontrar la solución particular, se toma la elección fácil, es decir, $s_{r+1} = s_{r+2} = \cdots = s_n = 0$. Este paso es muy simple, ya que basta resolver el sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r = b_1 \\ \qquad\quad a_{22}x_2 + \cdots + a_{2r}x_r = b_2 \\ \qquad\qquad\qquad\qquad\quad \vdots \\ \qquad\qquad\qquad\qquad a_{rr}x_r = b_r, \end{cases}$$

que es muy fácil de resolver, ya sea directamente por sustitución, o usando la regla de Cramer. A esta solución particular encontrada del sistema (5.8) la denotamos por X_0.

Paso 4. Finalmente, en virtud del Teorema 5.5.1 se encuentran todas las soluciones. Para esto, basta encontrar una base del subespacio de soluciones del sistema homogéneo asociado a (5.8)

Esta base se obtiene directamente aplicando de nuevo el Teorema 5.4.2, $n - r$ veces, ahora al sistema homogéneo asociado para los valores

$$\begin{cases} s_{r+1} = 1, & s_{r+2} = 0, & \cdots, & s_n = 0 \\ s_{r+1} = 0, & s_{r+2} = 1, & \cdots, & s_n = 0 \\ \quad\vdots & \quad\vdots & \ddots & \quad\vdots \\ s_{r+1} = 0, & s_{r+2} = 0, & \cdots, & s_n = 1. \end{cases}$$

Las soluciones se obtienen fácilmente, por sustitución, o usando la regla de Cramer.

Los vectores solución en $\mathbb{R}^n$ generados por estos valores que denotamos por $X_1, \ldots, X_{n-r}$ son evidentemente linealmente independientes (ya que tienen una submatriz de $n - r \times n - r$ con determinante igual a 1), y por lo tanto forman una base. Esto último se sigue, ya que si se toma $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ determinada por

$$x \longmapsto x_1 C_1 + \cdots + x_n C_n,$$

donde las C_i son las columnas de (5.8), se tiene en virtud del Teorema 5.4.4 que la dimensión del *kernel* de f es $n - r$.

Por consiguiente, todas las soluciones de (5.8) *están dadas por*

$$\left\{ X_0 + \sum_{i=1}^{n-r} t_i X_i \,|\, t_i \in \mathbb{R},\ i \in \{1, 2, ..., n-r\} \right\}.$$

Ejemplos

1. Para resolver el sistema

$$\left\{\begin{array}{rcrcrcrcl} x & + & y & - & 2z & & & = & 1 \\ 2x & - & y & - & z & - & 3t & = & -4 \\ x & + & 2y & - & 3z & + & t & = & 3 \\ 2x & + & y & - & 3z & - & t & = & 0, \end{array}\right.$$

se aplican operaciones elementales a la matriz del sistema

$$\left(\begin{array}{rrrrr} 1 & 1 & -2 & 0 & 1 \\ 2 & -1 & -1 & -3 & -4 \\ 1 & 2 & -3 & 1 & 3 \\ 2 & 1 & -3 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{rrrrr} 1 & 1 & -2 & 0 & 1 \\ 0 & -3 & 3 & -3 & -6 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & -1 & 1 & -1 & -2 \end{array}\right)$$

$$\sim \left(\begin{array}{rrrrr} 1 & 1 & -2 & 0 & 1 \\ 0 & -1 & 1 & -1 & -2 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{rrrrr} 1 & 1 & -2 & 0 & 1 \\ 0 & -1 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Se concluye que el sistema original tiene las mismas soluciones que el sistema

$$\left\{\begin{array}{rcrcrcrcl} x & + & y & - & 2z & & & = & 1 \\ & - & y & + & z & - & t & = & -2. \end{array}\right.$$

Tomando los valores $(t, z) = (0, 0)$, se obtiene la solución particular $X_0 = (-1, 2, 0, 0)$.

Resolvemos ahora el sistema homogéneo asociado. Estas mismas operaciones elementales por renglones se aplican para simplificar el sistema homogéneo asociado, la última columna, que son ceros, no cambia. Nótese que la dimensión del subespacio vectorial de las soluciones es $4 - 2 = 2$. Se puede encontrar una base de éste tomando para (z, t), los valores $(1, 0)$ y $(0, 1)$.

Si $z = 1$ y $t = 0$ se tiene

$$\left\{\begin{array}{rcrcrcl} x & + & y & - & 2 & = & 0 \\ & - & y & + & 1 & = & 0, \end{array}\right.$$

cuya solución es $X_1 = (1, 1, 1, 0)$. En el caso de $z = 0$, y $t = 1$ se tiene

$$\begin{cases} x + y = 0 \\ - y + -1 = 0, \end{cases}$$

i.e., $X_2 = (1, -1, 0, 1)$. Por lo cual el conjunto de soluciones del sistema está dado por

$$\{X_0 + t_1 X_1 + t_2 X_2 \mid t_1, t_2 \in \mathbb{R}\},$$

o alternativamente, usando las variables originales

$$\{X_0 + z X_1 + t X_2 \mid z, t \in \mathbb{R}\}.$$

2.

$$\begin{cases} x - y + z + 2t = 0 \\ 2x + y - z - 5t = -6 \\ x - 2y + z + 4t = -1 \\ x + y + 2z - t = 5. \end{cases}$$

La matriz aumentada del sistema es

$$\begin{pmatrix} 1 & -1 & 1 & 2 & 0 \\ 2 & 1 & -1 & -5 & -6 \\ 1 & -2 & 1 & 4 & -1 \\ 1 & 1 & 2 & -1 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 & 2 & 0 \\ 0 & 3 & -3 & -9 & -6 \\ 0 & -1 & 0 & 2 & -1 \\ 0 & 2 & 1 & -3 & 5 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & -1 & 1 & 2 & 0 \\ 0 & 1 & -1 & -3 & -2 \\ 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 & 2 & 0 \\ 0 & 1 & -1 & -3 & -2 \\ 0 & 0 & -1 & -1 & -3 \\ 0 & 0 & 1 & 1 & 3 \end{pmatrix},$$

y el sistema es equivalente a

$$\begin{cases} x - y + z + 2t = 0 \\ y - z - 3t = -2 \\ z + t = 3. \end{cases}$$

Tomando el valor $t = 0$, se obtiene de manera inmediata una solución particular $X_0 = (-2, 1, 3, 0)$ (se encuentra z, luego y y finalmente x).

Ahora resolvemos el sistema homogéneo asociado, tomando $t = 1$, se tiene

$$\begin{cases} x - y + z = -2 \\ \quad\; y - z = 3 \\ \quad\quad\; z = -1, \end{cases}$$

cuya solución es $X_1 = (1, 2, -1, 1)$. En consecuencia, el conjunto de soluciones está dado por

$$\{X_0 + t\, X_1 \;\; | \;\; t \in \mathbb{R}\}.$$

Nótese que el conjunto de soluciones se puede expresar en términos de las variables originales, como en los dos ejemplo anteriores. Esto siempre se puede hacer, ya que las variables *libres* pueden ser usadas como parámetros, ya que los vectores elegidos tienen coordenadas cero en estas variables, salvo en una de ellas donde toman el valor 1.

3.

$$\begin{cases} x - y + 2z = 1 \\ \quad\; y - z = 1 \\ 3x + y - z = 0 \\ 4x + y \quad\; = 2. \end{cases}$$

La matriz aumentada del sistema está dada por

$$\begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 3 & 1 & -1 & 0 \\ 4 & 1 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 4 & -7 & -3 \\ 0 & 5 & -8 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -3 & -7 \\ 0 & 0 & -3 & -7 \end{pmatrix},$$

i.e., las soluciones son las mismas que las del sistema

$$\begin{cases} x - y + 2z = 1 \\ \quad\; y - z = 1 \\ \quad\quad\; 3z = 7. \end{cases}$$

Por consiguiente, la única solución es

$$X_0 = \left(\frac{-1}{3}, \frac{10}{3}, \frac{7}{3}\right), \quad \text{ya que} \quad x = \frac{3}{3} - \frac{14}{3} + \frac{10}{3}.$$

4. Considérese el sistema

$$\begin{cases} x_1 + x_2 + 2\,x_4 = 2 \\ x_2 + 2\,x_3 = 0 \\ x_4 = 2. \end{cases}$$

La matriz aumentada asociada es

$$\begin{pmatrix} 1 & 1 & 0 & 2 & 2 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

Aunque esta matriz es escalonada, las primeras tres columnas no son linealmente independientes, por lo que se debe renombrar las variables.

Sean $y_1 = x_1,\ y_2 = x_2,\ y_3 = x_4$ y $y_4 = x_3,$ entonces el sistema se escribe

$$\begin{cases} y_1 + y_2 + 2\,y_3 = 2 \\ y_2 + 2\,y_4 = 0 \\ y_3 = 2, \end{cases}$$

y la matriz aumentada del nuevo sistema es

$$\begin{pmatrix} 1 & 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{pmatrix}.$$

Obsérvese que esta matriz se obtuvo al intercambiar dos columnas de la original. Ahora podemos encontrar las soluciones, tomando $y_4 = 0$, se obtiene una solución particular

$$Y_0 = (-2, 0, 2, 0).$$

Para el sistema homogéneo, si $y_4 = 1$ se obtiene la solución

$$Y_1 = (2, -2, 0, 1),$$

y todas las soluciones son

$$\{Y_0 + y_4\,Y_1 \ \mid \ y_4 \in \mathbb{R}\}.$$

En términos de las variables originales se obtiene

$$X_0 = (-2, 0, 0, 2), \qquad X_1 = (2, -2, 1, 0),$$

y todas las soluciones son

$$\{X_0 + x_3\, X_1 \;\mid\; x_3 \in \mathbb{R}\}.$$

EJERCICIO 5.6

1. Resolver los siguientes sistemas de ecuaciones encontrando una solución particular y una base del subespacio de soluciones del sistema homogéneo asociado.

$$a)\begin{cases} 2x + 3y + z - t + w = 2 \\ 3x + z + 2t - 5\,w = 0 \\ x + y - 2z - 4\,w = 0, \end{cases}$$

$$b)\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 7 \\ 2\,x_2 + x_3 - 4\,x_5 = 8 \\ 4\,x_1 + 2\,x_5 - x_6 = 2, \end{cases}$$

$$c)\begin{cases} x + 3y + z + t - w = 2 \\ 3\,y - 2\,z - 3\,w = 0 \\ 2\,t + 5\,w = 2, \end{cases}$$

$$d)\begin{cases} x - 2y + z + t = 5 \\ y + 2\,z - t = 3 \\ 5\,z - 2\,t = -1. \end{cases}$$

Capítulo 6

Los anillos $\mathbb{Z}$ y $\mathbb{Z}_m$

Se establecen las propiedades del anillo de los enteros de manera axiomática, se definen los anillos $\mathbb{Z}_m$, y se prueba que en efecto son anillos. Posteriormente, se discuten los dominios enteros, el orden y las unidades. El libro concluye probando la equivalencia del principio de inducción con el del buen orden.

6.1. Anillos

Definición 62 *El conjunto de los enteros, denotados por* $\mathbb{Z}$, *consiste de los números*

$$\{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

Un axioma es un fundamento de verdad o un hecho evidente que se acepta por verdadero pero no se demuestra. Los axiomas que fundamentan la matemática son pocos, éstos esencialmente son los de Peano, que están basados en razonamientos que de alguna manera, su validez es evidente, dicho de otra manera, se aceptan como verdaderos. El estudio de estos axiomas son objeto de estudio en los cursos de Lógica Matemática y Teoría de Conjuntos. Recordamos ahora la siguiente importante definición.

Definición 63 *Sea* A *un conjunto, una operación binaria en* A *es una función*

$$\mu : A \times A \longrightarrow A.$$

Por ejemplo, la suma y la multiplicación en $\mathbb{Z}$ son operaciones binarias. Recordamos del primer capítulo que éstas cumplen los siguientes axiomas.

S1) $a + b = b + a \quad \forall\, a, b \in \mathbb{Z}$ (conmutatividad),

S2) $(a + b) + c = a + (b + c) \quad \forall\, a, b, c \in \mathbb{Z}$ (asociatividad),

S3) $\exists\, 0 \in \mathbb{Z} \ni a + 0 = a \quad \forall\, a \in \mathbb{Z}$ (existencia del neutro aditivo),

S4) $a + (-a) = 0 \quad \forall\, a \in \mathbb{Z}$ (existencia del inverso aditivo).

P1) $a\,b = b\,a \quad \forall\, a, b \in \mathbb{Z}$ (conmutatividad),

P2) $(a\,b)\,c = a\,(b\,c) \quad \forall\, a, b, c \in \mathbb{Z}$ (asociatividad),

P3) $\exists\, 1 \in \mathbb{Z} \ni a \cdot 1 = a \quad \forall\, a \in \mathbb{Z}$ (existencia del neutro multiplicativo),

D) $a\,(b + c) = a\,b + a\,c \quad \forall\, a, b, c \in \mathbb{Z}$ (distributividad).

Definición 64 *A un conjunto X con dos operaciones binarias que satisfacen los axiomas $S1, S2, S3, S4, P1, P2, P3$ y D se le llama anillo conmutativo con elemento unitario (o unidad). Si no satisface $P1$ y $P3$ se le llama simplemente anillo.*

Con esta terminología, $\mathbb{Z}$ es un anillo conmutativo con unidad. No se prueba este hecho, aceptaremos las propiedades $S1, S2, S3, S4, P1, P2, P3$ y D como ciertas, y simplemente nos referiremos a éstas como los primeros ocho axiomas de los enteros, llamándolos **Axioma 1**, **Axioma 2**,..., y **Axioma 8**, respectivamente. Estos axiomas, junto con otros axiomas, se tomarán como el fundamento para el estudio de los anillos $\mathbb{Z}_m$ y de otros temas que se estudian en el segundo curso Álgebra Superior II, como por ejemplo, la divisibilidad y los números reales.

6.2. Anillos $\mathbb{Z}_m$

Sea $A = \{\overline{0}, \overline{1}\}$, se definen operaciones de suma y producto mediante las siguientes tablas.

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$\times$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

Figura 6.1: Tablas de suma y producto en $\mathbb{Z}_2$

Se les puede llamar suma y producto a estas operaciones, el conjunto A con estas operaciones resulta ser un anillo conmutativo con unidad, que se le denota por $\mathbb{Z}_2$.

Se deben verificar las propiedades de anillo, por ejemplo

$$\overline{0} + \overline{1} = \overline{1} + \overline{0}.$$

Demostrar otras propiedades es largo y tedioso. En particular, para probar la asociatividad hay que considerar todas las posibles ternas en los valores $\overline{0}$ y $\overline{1}$ (el número de posibilidades es $OR_2^3 = 2^3$). Más aún, hay ocho propiedades de anillo por probar. Una manera más eficaz es establecer un patrón general que ahora presentamos.

Recordamos del primer capítulo que dado un número $m \in \mathbb{N}$ fijo, la relación $a \sim b$, si $a - b = km$, $k \in \mathbb{Z}$ es de equivalencia. Denotaremos por $\overline{a}$ la clase de equivalencia a la que pertenece a.

Por ejemplo, si $m = 2$

$$\begin{aligned} \overline{0} &= \{\ldots, -4, -2, 0, 2, 4, \ldots\} = \{\textit{pares}\}, \\ \overline{1} &= \{\ldots, -3, -1, 1, 3, 5, \ldots\} = \{\textit{impares}\}. \end{aligned}$$

Esto se sigue, ya que

$$\overline{0} = \{t \in \mathbb{Z} \,|\, t \sim 0\} \quad \text{y} \quad \overline{1} = \{t \in \mathbb{Z} \,|\, t \sim 1\},$$

i.e., $t \sim 0$, si t es par. Y también $t \sim 1$, si $t - 1$ es par, lo cual sucede si t es impar.

Para el caso $m = 3$, las clases de equivalencia son

$$\begin{aligned} \overline{0} &= \{\ldots, -6, -3, 0, 3, 6, \ldots\}, \\ \overline{1} &= \{\ldots, -5, -2, 1, 4, \ldots\}, \\ \overline{2} &= \{\ldots, -4, -1, 2, 5, \ldots\}. \end{aligned}$$

Una mejor descripción sería

$$\begin{aligned} \overline{0} &= \{t \mid t = 3k, \quad \textit{i.e., } t \textit{ es un múltiplo de } 3\}, \\ \overline{1} &= \{t \mid t - 1 = 3k, \quad \textit{i.e., } t - 1 \textit{ es un múltiplo de } 3\}, \\ \overline{2} &= \{t \mid t - 2 = 3k, \quad \textit{i.e., } t - 2 \textit{ es un multiplo de } 3\}. \end{aligned}$$

Obsérvese que con la convención adaptada se tiene que en el presente caso $m = 3$, por ejemplo, $\overline{5} = \overline{2} = \overline{-1}$.

Definición 65 *Al conjunto de clases de equivalencia en* $\mathbb{Z}$ *definidas por la relación*

$$a \sim b \quad si \quad a - b = km, \quad k \in \mathbb{Z}, \quad m \in \mathbb{N} \ \textit{fijo},$$

se le denota por $\mathbb{Z}_m$.

Resulta que este conjunto es un anillo conmutativo con unidad, lo cual probamos ahora. Obsérvese primero que

$$\mathbb{Z}_m = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{m-1}\}$$

(ejercicio).

Se definen dos operaciones en $\mathbb{Z}_m$

$$\overline{a} + \overline{b} = \overline{a_1 + b_1}, \ \text{ donde } \ a_1 \in \overline{a} \ \text{ y } \ b_1 \in \overline{b},$$

$$\overline{a}\,\overline{b} = \overline{a_1\, b_1}, \ \text{ donde } \ a_1 \in \overline{a} \ \text{ y } \ b_1 \in \overline{b}.$$

Hay que probar que estas operaciones están bien definidas y no dependen de los representantes a_1 y b_1. Si

$$a_1 \sim a_2 \quad \text{y} \quad b_1 \sim b_2,$$

se tiene

$$a_1 - a_2 = m\,k_1 \qquad \text{y} \qquad b_1 - b_2 = m\,k_2, \quad \text{donde} \quad k_1,\, k_2 \in \mathbb{Z}.$$

Por lo tanto

$$a_1 + b_1 - (a_2 + b_2) = m(k_1 + k_2)$$
$$\therefore \quad a_1 + b_1 \sim a_2 + b_2.$$

También,

$$\begin{aligned} a_1\, b_1 - a_2\, b_2 &= a_1\, b_1 - a_2\, b_1 + a_2\, b_1 - a_2\, b_2 \\ &= (a_1 - a_2)\, b_1 + a_2\, (b_1 - b_2) \\ &= m\, k_1\, b_1 + a_2\, m\, k_2 = m(\textit{entero}), \end{aligned}$$

y $a_1\, b_1 \sim a_2\, b_2$.

Resulta ahora fácil probar para todos los conjuntos $\mathbb{Z}_m$ que son en efecto anillos conmutativos con unidad.

Teorema 6.2.1 $\forall\, m \in \mathbb{N}$, $\mathbb{Z}_m$ *es un anillo conmutativo con unidad.*

Demostración. El teorema se sigue esencialmente del hecho de que estas mismas propiedades las cumplen los enteros y de que las operaciones están bien definidas. Probamos seis de estas propiedades y las dos restantes quedan como ejercicio. Tomando cualesquiera representantes $a, b,$ y c de $\overline{a}$, $\overline{b}$ y $\overline{c}$, se tiene

S1) $\overline{a} + \overline{b} = \overline{a+b} = \overline{b+a} = \overline{b} + \overline{a}$,

S2) $\overline{a} + (\overline{b} + \overline{c}) = \overline{a} + \overline{b+c} = \overline{a+(b+c)} = \overline{(a+b)+c} = \overline{a+b} + \overline{c}$
$= (\overline{a} + \overline{b}) + \overline{c}$,

S3) $\overline{a} + \overline{0} = \overline{a+0} = \overline{a}$,

P1) $\overline{a} \cdot \overline{1} = \overline{a \cdot 1} = \overline{a}$,

P3) $\overline{a}\,\overline{b} = \overline{ab} = \overline{ba} = \overline{b}\,\overline{a}$,

D) $\overline{a}\,(\overline{b} + \overline{c}) = \overline{a}(\overline{b+c}) = \overline{a(b+c)} = \overline{ab+ac} = \overline{ab} + \overline{ac} = \overline{a}\,\overline{b} + \overline{a}\,\overline{c}$.

□

Habiendo probado que $\mathbb{Z}_m$ es un anillo podemos usar m símbolos únicos para los elementos. Por ejemplo, $\mathbb{Z}_3 = \{\overline{0}, \overline{1}, \overline{2}\}$. Sin embargo, esto no impide pensar, o escribir, cualquier elemento en otros términos, por ejemplo, en $\mathbb{Z}_3$ $\overline{4}$ es $\overline{7}$. Esta manera de razonar es la herramienta adecuada para sumar y multiplicar.

La relación $a \sim b$ se escribirá posteriormente como $a \equiv b \ mod\ m$, se dirá que *a es congruente con b módulo m.* Este tema se estudiará con más detalle en el segundo curso.

Podemos ahora, sin mayor dificultad, establecer las tablas de multiplicar de algunos de estos anillos, por ejemplo $\mathbb{Z}_4$.

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

$\times$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Figura 6.2: Tablas de suma y producto en $\mathbb{Z}_4$

Nótese la simetría en las tablas, la cual se debe a la conmutatividad de la suma y del producto.

EJERCICIOS 6.2
1. Sea $m \in \mathbb{N}$ fijo. Demuestre que $\mathbb{Z}_m = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{m-1}\}$. Sugerencia: usar el algoritmo de la división.
2. Terminar la prueba de que $\mathbb{Z}_m$ es una anillo conmutativo con unidad, es decir, hay que probar las propiedades S4) y P2).
3. Escribir las tablas de $\mathbb{Z}_3$, $\mathbb{Z}_5$ y $\mathbb{Z}_6$.

6.3. Propiedades de los enteros

Probaremos primero ciertas propiedades que sólo dependen de los axiomas de la suma y el producto junto con la distributividad. Por lo que son también válidas para todo anillo comnutativo con unidad.

LEY DE LA CANCELACIÓN POR LA IZQUIERDA

Dados $a, b, c \in \mathbb{Z}$ *tales que* $a + b = a + c$, *se tiene* $b = c$.

Usando la existencia del inverso y la asociatividad se tiene

$$(-a) + (a + b) = (-a) + (a + c) \quad \text{y} \quad b = c.$$

LEY DE LA CANCELACIÓN POR LA DERECHA

Dados $a, b, c \in \mathbb{Z}$ *tales que* $a + c = b + c$, *se tiene* $a = b$.

Esta ley se sigue ahora de la conmutatividad. Como consecuencia se tiene que el neutro aditivo es único. Si $a+b = a$, entonces $a+b = a+0$, y $b = 0$.

Proposición 6.3.1 $\forall\, a \in \mathbb{Z}$, $a \cdot 0 = 0$.

DEMOSTRACIÓN. Como

$$0 + 0 \cdot a = 0 \cdot a = (0 + 0)\, a = 0 \cdot a + 0 \cdot a,$$

por cancelación

$$0 = 0 \cdot a.$$

□

La cancelación implica también que el inverso aditivo es único. Si $a+c = 0$, entonces $a + (-a) = a + c$, por lo tanto $-a = c$.

Proposición 6.3.2 $\forall\, a \in \mathbb{Z} \quad -(-a) = a.$

Demostración. Como $a + (-a) = 0,$ entonces el inverso aditivo de $(-a),$ es decir, $-(-a)$ es $a,$ por unicidad.

□

Una propiedad fundamental es la siguiente.

Proposición 6.3.3 *(Regla de los signos)* $\forall\, a, b \in \mathbb{Z}$ *se tiene*

i) $(-a)\, b = -(a\, b) = a\, (-b),$

ii) $(-a)(-b) = a\, b.$

Demostración. Como

$$(-a)\, b + a\, b = ((-a) + a)\, b = 0,$$

por unicidad

$$(-a)\, b = -(a\, b).$$

La otra igualdad en $i)$ se prueba de manera análoga.

Para probar $ii),$ usando $i)$ se tiene

$$(-a)(-b) + [-(ab)] = (-a)(-b) + [(-a)\, b] = (-a)(-b + b) = 0,$$

por lo que (por unicidad) $(-a)(-b) = a\, b.$

□

Como caso particular de la Proposición 6.3.3 se tiene $(-1)\, a = -a$ y $(-1)(-1) = 1.$

Definición 66 *Sean* $a, b \in \mathbb{Z},$ *se define la resta o diferencia de* a *menos* b *como*

$$a + (-b),$$

se denota por $a - b.$

Obsérvese que $a\, (b - c) = ab - ac,$ ya que

$$a(b + (-c)) = a\, b + a\, (-c) = a\, b + (-(ac)),$$

por la regla de los signos.

Nótese que la unicidad del inverso aditivo implica que

$$-(a+b) = -a - b,$$

ya que ambos son inversos aditivos de $a+b$.

Los enteros tienen otras propiedades que no son consecuencia de los primeros 8 axiomas: S1), S2), S3), S4), P1), P2), P3) y D), que los hemos denotado también como **Axioma 1, Axioma 2,**..., **Axioma 8**.

Axioma 9 $\forall\ a, b \in \mathbb{Z} - \{0\}$, *se tiene* $a\,b \neq 0$.

Una formulación equivalente es decir $\forall\ a, b \in \mathbb{Z}$ ∍ $a\,b = 0$ se tiene que $a = 0$ o $b = 0$. Esta propiedad válida en $\mathbb{Z}$ no se cumple en otros anillos conmutativos con unidad, por ejemplo en $\mathbb{Z}_4$, $\overline{2} \cdot \overline{2} = \overline{0}$, o en $\mathbb{Z}_9$, $\overline{3} \cdot \overline{3} = \overline{0}$. También, en $\mathbb{Z}_6$, $\overline{2} \cdot \overline{3} = \overline{0}$.

Obsérvese que la existencia de estos ejemplos prueba que, en efecto, el **Axioma 9** no es consecuencia de los primeros 8 axiomas.

Una manera intuitiva de convencerse de la *veracidad* del **Axioma 9** es suponer $a\,b = 0$, introducir sus inversos multiplicativos, es decir, los números racionales $1/a$ y $1/b$ y llegar a una contradicción:

$$1 = \left(\frac{1}{a}\right)\left(\frac{1}{b}\right) a\,b = \left(\frac{1}{a}\right)\left(\frac{1}{b}\right) 0 = 0.$$

Sin embargo, este argumento no tiene validez formal, ya que no se han definido los números racionales.

Definición 67 *Sean A un anillo, y $a, b \in A$ tales que $a\,b = 0$, entonces a los elementos a y b se les llama divisores del cero.*

El axioma 9 dice que en $\mathbb{Z}$ no hay divisores de cero distintos de cero.

Definición 68 *Un dominio entero es un anillo conmutativo con unidad que no tiene divisores de 0 distintos de 0.*

Un ejemplo de dominio entero es $\mathbb{Z}$, mientras que $\mathbb{Z}_4$, $\mathbb{Z}_6$, $\mathbb{Z}_9$, $\mathbb{Z}_{18}$, no lo son. Obsérvese que en un dominio entero A vale la LEY DE LA CANCELACIÓN PARA LA MULTIPLICACIÓN:

$$\textit{Sean } a, b, c \in A,\ \ a \neq 0,\ \ \textit{tales que } a\,b = a\,c,\ \ \ \textit{entonces } b = c.$$

Esta ley se sigue ya que la hipótesis implica $a\,b - a\,c = 0$, por lo cual se tiene $a\,(b-c) = 0$, y dado que $a \neq 0$, necesariamente $b - c = 0$.

En particular, esta ley es válida en $\mathbb{Z}$. Nótese que es necesaria la condición $a \neq 0$, ya que $0 \cdot b = 0 \cdot c \ \ \forall\, b, c \in A$. El recíproco es cierto: si A es un anillo conmutativo con 1, para el cual vale la ley de la cancelación para el producto, entonces A es un dominio entero (ejercicio).

EJERCICIO 6.3
1. Probar que si A es un anillo conmutativo con 1, para el cual vale la ley de la cancelación para el producto, entonces A es un dominio entero.

6.4. Orden y unidades en $\mathbb{Z}$

Se quiere formalizar el hecho de que un número entero es mayor que otro si está más a la derecha en la recta.

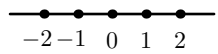


Figura 6.3: Los enteros como puntos en la recta real

Este concepto lo podemos relacionar con los naturales, $\mathbb{N} = \{1, 2, 3, \ldots\}$, primero necesitamos algunos axiomas de cerradura y tricotomía.

Axioma 10 *Si* $a, b \in \mathbb{N}$, *entonces* $a + b \in \mathbb{N}$.

Axioma 11 *Si* $a, b \in \mathbb{N}$, *entonces* $a\,b \in \mathbb{N}$.

Axioma 12 *Si* $a \in \mathbb{Z}$, *entonces se cumple una y sólo una de las siguientes afirmaciones*:

i) $a \in \mathbb{N}$,

ii) $a = 0$,

iii) $-a \in \mathbb{N}$.

Estos axiomas junto con las leyes de los signos establece que $\mathbb{Z}$ es un conjunto cerrado bajo la suma y el producto. Como se mencionó antes, estos 12 axiomas de los enteros se basan en otro conjunto menor de axiomas llamados de Peano, y éstos a su vez se basan en otros pocos axiomas fundamentales que de alguna manera son hechos contundentes, por ejemplo, el axioma de extensión que establece que dos conjuntos son iguales si tienen los mismo elementos, otros axiomas son el de especificación y el de infinito. Cf. [3]

Definición 69 *Sean $a, b \in \mathbb{Z}$, se dice que a es mayor que b, si*

$$a - b \in \mathbb{N},$$

Se escribe $a > b$.

Obsérvese que

$$a > 0 \quad \Leftrightarrow \quad a - 0 \in \mathbb{N} \quad \Leftrightarrow \quad a \in \mathbb{N}.$$

Proposición 6.4.1 *(Transitividad) Sean $a, b, c \in \mathbb{Z}$ tales que $a > b$ y $b > c$, entonces $a > c$.*

Demostración. Como $a - b \in \mathbb{N}$, y $b - c \in \mathbb{N}$, se tiene

$$a - b + b - c = a - c \in \mathbb{N}$$

(en virtud del Axioma 10), es decir $a > c$.

□

Obsérvese que el **Axioma 12** se puede reescribir de la siguiente manera: dado $a \in \mathbb{Z}$, sucede una y sólo una de las siguientes tres afirmaciones:

$$a > 0, \quad a < 0, \quad \text{o} \quad a = 0.$$

Esto se sigue ya que $a \in \mathbb{N} \quad \Leftrightarrow \quad a > 0$,

$$a < 0 \quad \Leftrightarrow \quad 0 > a \quad \Leftrightarrow \quad 0 - a \in \mathbb{N} \quad \Leftrightarrow \quad -a \in \mathbb{N}.$$

El orden es compatible con la suma y el producto.

Proposición 6.4.2 *Sean $a, b, c \in \mathbb{Z}$ tales que $a > b$, entonces*

$$a + c > b + c.$$

Demostración. Como $a - b \in \mathbb{N}$, se tiene

$$a + c - (b + c) \in \mathbb{N}.$$

□

Proposición 6.4.3 *Sean $a, b, c \in \mathbb{Z}$ tales que $a > b$ y $c > 0$, entonces*

$$a\,c > b\,c.$$

Demostración. $a-b \in \mathbb{N}$ y $c \in \mathbb{N}$, por lo cual

$$(a-b)\,c = a\,c - b\,c \in \mathbb{N},$$

es decir, $a\,c > b\,c$.

□

otras propiedades de orden

1. $a^2 > 0 \quad \forall\, a \in \mathbb{Z} - \{0\}$:

   Si $a \in \mathbb{N}$, se sigue del **Axioma 11** que también a^2 es un natural. Si $-a \in \mathbb{N}$, $(-a)(-a) \in \mathbb{N}$ y $a^2 \in \mathbb{N}$, por la regla de los signos.

2. *Si* $a \geq 0$ y $b \geq 0$, *entonces* $a+b \geq 0$:

   $a, b \in \mathbb{N} \quad \Rightarrow \quad a+b \in \mathbb{N}.$

3. $a < b \quad \Leftrightarrow \quad -a > -b \quad (-b < -a)$:

   $b-a \in \mathbb{N} \quad \Leftrightarrow \quad -a-(-b) \in \mathbb{N}.$

unidades en $\mathbb{Z}$

Se mostró que todos los elementos de $\mathbb{Z}$ tienen un inverso aditivo, sin embargo, sólo 1 y -1 tienen inverso multiplicativo. Lo cual se prueba a continuación.

Definición 70 *En un anillo a los elementos que tienen inverso multiplicativo se les llama unidades.*

Por ejemplo, en $\mathbb{Z}_5$ todos los elementos no cero son unidades, ya que $\overline{2} \cdot \overline{3} = \overline{1}$, y $\overline{4} \cdot \overline{4} = 1$, sin embargo, en $\mathbb{Z}_6$, $\overline{2}$ no es unidad, ya que

$$\overline{2} \cdot \overline{1} = \overline{2}, \ \ \overline{2} \cdot \overline{2} = \overline{4}, \ \ \overline{2} \cdot \overline{3} = \overline{0}, \ \ \overline{2} \cdot \overline{4} = \overline{2}, \ \ \overline{2} \cdot \overline{5} = \overline{4}, \ \ \overline{2} \cdot \overline{0} = \overline{0}.$$

Proposición 6.4.4 *Las únicas unidades en* $\mathbb{Z}$ *son 1 y -1.*

Demostración. 0 no es unidad, ya que $0\,a = 0 \quad \forall\, a \in \mathbb{Z}$. Como $1 \cdot 1 = 1$ y $(-1)(-1) = 1$, 1 y -1 sí lo son.

Ahora, si $a > 1$, a no es unidad. Probamos que $a\,b \neq 1 \ \ \forall\, b \in \mathbb{Z}$. Primero, $a\,0 = 0$ y $a\,1 = a$. Si $b < 0$, entonces $a\,b < a \cdot 0 = 0$ y $a\,b \neq 1$. También si $b > 1$, $a\,b > a\,1 = a$ y $a\,b \neq 1$.

Si $a < -1$, a tampoco es unidad. Si a fuera unidad, se tendría $a\,b = 1$ y $(-a)(-b) = 1$, *i.e.*, $-a$, que es mayor a 1, sería unidad, contradiciendo la 1a parte.

□

EJERCICIO 6.4
1. Encuentre las unidades en los anillos $\mathbb{Z}_4$, $\mathbb{Z}_6$, y $\mathbb{Z}_7$.

6.5. Principio de inducción

Profundizamos sobre este tema que se presentó al inicio del libro.

Axioma 13 (Principio de inducción)

Sea M un subconjunto de los naturales tales que se cumplen las siguientes condiciones:

i) $1 \in M$

ii) *si* $n \in M$, *entonces* $n+1 \in M$.

Bajo estas hipótesis se concluye que

$$M = \mathbb{N}.$$

Este axioma, o principio, dice que si un subconjunto de los naturales cumple las condiciones *i)* y *ii)*, entonces ese subconjunto consiste de todos los naturales.

Su importancia radica en ser muy útil para probar propiedades que cumplen todos los naturales, por ejemplo, para probar la identidad

$$1 + 3 + 5 + 7 + \cdots + (2n-1) = n^2. \tag{6.1}$$

Sea $M \subset \mathbb{N}$, aquellos números que cumplen (6.1), se tiene que

i) $1 \in M$, ya que $1 = 1^2$,

ii) si $n \in M$, entonces

$$1 + 3 + 5 + 7 + \cdots + (2n-1) + (2n+1) = n^2 + (2n+1) = (n+1)^2,$$

por lo que $n+1 \in M$.

El principio de inducción implica que (6.1) es válido para todos los naturales.

Como un segundo ejemplo, considérese la igualdad

$$2^0 + 2^1 + 2^2 + \cdots + 2^{n-1} = 2^n - 1. \tag{6.2}$$

La prueba de esta identidad se puede interpretar como sigue: para cada número n, la identidad (6.2) es una proposición que denotamos por P_n, si se logra probar que P_1 es correcta, y que cada vez que P_n lo es, la proposición P_{n+1} también lo es, se concluye que (6.2) se cumple $\forall\, n \in \mathbb{N}$.

i) $2^{1-1} = 2^0 = 2^1 - 1$ y por lo tanto P_1 se cumple,

ii) si P_n se cumple

$$2^0 + 2^1 + 2^2 + \cdots + 2^{n-1} + 2^n = 2^n - 1 + 2^n = 2 \cdot 2^n - 1 = 2^{n+1} - 1,$$

por lo que P_{n+1} se cumple.

Algunas veces es útil usar un principio equivalente al de inducción.

PRINCIPIO DE INDUCCIÓN MODIFICADO

Sea $M \subset \mathbb{N}$ *tal que cumple las siguientes dos condiciones*:

a) $1 \in M$

b) *cada vez que* $1, 2, \ldots, n \in M$, *se tiene que* $n + 1 \in M$.

Bajo estas hipótesis se concluye que $M = \mathbb{N}$.

Teorema 6.5.1 *El principio de inducción es equivalente al principio de inducción modificado.*

Demostración. Denotamos por ① al principio de inducción y por ② al principio modificado. Para probar que ① implica ②, hay que probar que las hipótesis de ② implican las de ①. Para esto se debe probar que si $n \in M$, entonces $n+1 \in M$. Aplicando *a*) junto con n veces la hipótesis *b*) se tiene que si $n \in M$, entonces $n + 1 \in M$: $1, 2 \in M \Rightarrow 3 \in M$, $1, 2, 3 \in M \Rightarrow 4 \in M$. Iterando este proceso se obtiene $n + 1 \in M$, por lo que se cumple *ii*). Por lo tanto $M = N$, y se sigue ②.

Finalmente para probar que ② implica ①, se debe probar que las hipótesis de ① implican las de ②, lo cual es evidente. □

El texto concluye estableciendo otra propiedad de los enteros que sorprendentemente es equivalente al principio de inducción.

PRINCIPIO DEL BUEN ORDEN

Sea $A \subset \mathbb{N},\ A \neq \emptyset$, *entonces* A *tiene un menor elemento, i.e.,* $\exists\, m_0 \in \mathbb{N}$, *tal que* $m_0 \leq n \quad \forall\, n \in A$.

Teorema 6.5.2 *El principio de inducción es equivalente al de buen orden.*

Demostración. Probamos primero que el de inducción implica el del buen orden. Si $A \subset \mathbb{N}$, $A \neq \emptyset$, y A no tiene un menor elemento, sea

$$B = \{m \in \mathbb{N} \mid m < n\ \forall n \in A\},$$

es decir, el conjunto de las cotas inferiores de A. Obsérvese que A y B no se intersecan, ya que ningún elemento es menor a sí mismo.

Resulta que $1 \in B$: si $1 \in A$, 1 es el menor elemento. También si $k \in B$, $k+1$ también está en B, ya que de otra manera $k+1$ no es menor que algún elemento de A. Esta última situación sólo acontece si $k+1 \in A$, puesto que se está suponiendo que todos los elementos de A son mayores a k. Sin embargo si $k+1 \in A$, $k+1$ sería el menor elemento de A.

El principio de inducción implica entonces que $B = \mathbb{N}$ y $A = \emptyset$, lo cual contradice las hipótesis, $\therefore$ A tiene un menor elemento.

Finalmente, probamos que el principio de buen orden implica el de inducción. Sea $M \subset \mathbb{N}$ tal que cumple las hipótesis del principio de inducción, probamos que $M^c = \emptyset$. Si $M^c \neq \emptyset$, sea m el menor elemento de M^c, entonces $m-1 \in M$ y por hipótesis de inducción $m \in M$, lo cual contradice $m \in M^c$, $\therefore$ $M^c = \emptyset$.

□

EJERCICIO 6.5

1. Probar la siguiente identidad

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \left(\frac{n\,(n+1)}{2}\right)^2.$$

Glosario de símbolos

C_n^k	número de combinaciones de n elementos tomados de k en k, coeficiente binomial
$\mathbb{C}$	los números complejos
$\supset$	contiene a
$\therefore$	de dónde, por lo tanto
$\sim$	equivalente a, relacionado con
$\in$	pertenece a, es elemento de
$\notin$	no pertenece a, no es elemento de
$\exists$	existe
$\nexists$	no existe
$\cap$	intersección (de conjuntos)
$\mathbb{N}$	los números naturales
$\neq$	diferente a, no igual a
O_n^m	número de ordenaciones de n elementos tomados de m en m
OR_n^m	número de ordenaciones con repetición de n elementos tomados de m en m
$\forall$	para todo
P_n	número de permutaciones de n elementos
$\mathbb{Q}$	los números racionales
$\mathbb{R}$	los números reales
$\subset$	subconjunto de
$\sum$	sumatoria

$\sum_{k=0}^{n} a_k$	$= a_0 + a_1 + a_2 + a_3 + \cdots + a_n$
$\iff$	si y sólo si
$\ni$	tal que
$\mid$	tal que, para especificar pertenencia a un conjunto
$\cup$	unión (de conjuntos)
$\emptyset$	el conjunto vacío
$\mathbb{Z}$	los números enteros

Bibliografía

[1] Bravo A., Rincón H. y Rincón C., *Álgebra superior*, México, Facultad de Ciencias, UNAM, 2008.

[2] Cárdenas H., Lluis E., Raggí F., Tomás F., *Álgebra superior*, México, Trillas, 1982.

[3] Halmos, P., *Naive Set Theory*, Nueva York, Springer-Verlag, 1974.

[4] Lang, S., *Linear Algebra*, Addison-Wesley, 1972.

[5] Lascurain, A., *Álgebra superior II* (en preparación).

Índice analítico

abcisa, 54
anillo, 30, 148
    conmutativo, 148
axioma
    de extensión, 155
    de infinito, 155
    especificación, 155
axiomas
    Peano, 147

base, 74
buen orden
    principio del, 160

campo, 30
cardinalidad, 17
clase de equivalencia, 26, 149
coeficientes binomiales, 47
combinación lineal, 69
combinaciones, 34, 46
conjunto
    de combinaciones lineales, 70
conjunto (s), 1
    complemento de, 5
    diferencia de, 7
    finito, 17
    infinito, 18
    intersección de, 4
    unión de, 3
    vacío, 2
corresponencia biunívoca, 43

Cramer
    regla de, 127

De Morgan
    leyes de, 6
dependencia lineal, 71
determinante, 93
    propiedades, 95
dominio entero, 154

escalares, 53
espacio vectorial
    $\mathbb{R}^n$, 60
    sobre los reales, 62
estructuras algebraicas, 28

factorial
    n-factorial, 21
función (es), 10, 36
    biyectiva, 16, 39
    característica, 48
    codominio de, 10
    composición de, 12
    dominio de, 10
    extensión de, 41
    imagen de, 11, 40, 135
    inversa derecha de, 14
    inversa izquierda de, 14
    invertible, 14
    inyectiva, 15, 39
    kernel de, 134

lineal, 134
multilineal, 113
suprayectiva, 16, 39

grupo, 30

inducción
principio modificado de, 159
principio de, 19, 158
inversión, 88

ley
de cancelación para el producto, 29
de cancelación para la suma, 29
de la cancelación, 152
de tricotomía, 30
ley distributiva, 4

matrices, 81
matriz
aumentada, 121
cuadrada, 82
del sistema, 121
determinante, 126
diagonal, 82
escalonada, 86
menores, 102
rango de, 83, 112, 123
transpuesta, 101
triangular, 82

números
complejos, 28
enteros, 28, 147
naturales, 28
racionales, 28
reales, 28
Newton
teorema del binomio, 21

operación binaria, 147
ordenaciones, 32, 43
con repetición, 31, 41
ordenada, 54

pareja ordenada, 8
partición, 24
Pascal
fórmula del triángulo de, 49
teorema de, 21
permutación
impar, 89
inversa, 91
par, 89
permutaciones, 33, 45, 88
producto
cartesiano, 8, 42, 53
de un escalar por un vector, 53
por escalares, 60
punto, 65

regla de correspondencia, 11
relación, 9
de equivalencia, 23, 149
dominio de, 10
imagen de, 10

signos
regla de los, 153
sistema
homogéneo, 122, 132
homogéneo asociado, 136
solución de, 123
subconjunto, 2
subespacio
generado, 70
vectorial, 70

vectorial de $\mathbb{R}^n$, 63
vectorial generado, 76
suma, 60
de vectores, 53

transposición, 89

unidades, 157

Van der Monde, 130
determinante tipo, 108
vectores, 53
linealmente dependientes, 71
linealmente independiantes, 72
ortogonales, 65

Algebra superior I
editado por la Facultad de Ciencias
de la Universidad Nacional Autónoma de México,
se terminó de imprimir el 14 de septiembre de 2017
en los talleres de Gráfica premier S. A. de C. V.
5 de febrero 2309. Col. San Jerónimo Chicahualco
C.P. 52170. Metepec. Estado de México.

El tiraje fue de 1000 ejemplares.

Está impreso en papel cultural de 90 g.
En su composición se utilizó tipografía Computer Modern de
11:13.5, 14:16 y 16:18 puntos de pica.

Tipo de impresión: offset.

El cuidado de la edición estuvo a cargo de
Patricia Magaña Rueda